



**QUEEN'S
UNIVERSITY
BELFAST**

Joint Information and Jamming Beamforming for Secrecy Rate Maximization in Cognitive Radio Networks

Nguyen, V. D., Duong, T. Q., Dobre, O. A., & Shin, O-S. (2016). Joint Information and Jamming Beamforming for Secrecy Rate Maximization in Cognitive Radio Networks. *IEEE Transactions on Information Forensics and Security*, 11(11), 2609-2623. <https://doi.org/10.1109/TIFS.2016.2594131>

Published in:

IEEE Transactions on Information Forensics and Security

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Joint Information and Jamming Beamforming for Secrecy Rate Maximization in Cognitive Radio Networks

Van-Dinh Nguyen, Trung Q. Duong, *Senior Member, IEEE*, Octavia A. Dobre, *Senior Member, IEEE*, and Oh-Soon Shin, *Member, IEEE*

Abstract—In this paper, we consider the secure beamforming design for an underlay cognitive radio multiple-input single-output broadcast channel in the presence of multiple passive eavesdroppers. Our goal is to design a jamming noise (JN) transmit strategy to maximize the secrecy rate of the secondary system. By utilizing the zero-forcing method to eliminate the interference caused by JN to the secondary user, we study the joint optimization of the information and JN beamforming for secrecy rate maximization of the secondary system while satisfying all the interference power constraints at the primary users, as well as the per-antenna power constraint at the secondary transmitter. For an optimal beamforming design, the original problem is a nonconvex program, which can be reformulated as a convex program by applying the rank relaxation method. To this end, we prove that the rank relaxation is tight and propose a barrier interior-point method to solve the resulting saddle point problem based on a duality result. To find the global optimal solution, we transform the considered problem into an unconstrained optimization problem. We then employ Broyden-Fletcher-Goldfarb-Shanno (BFGS) method to solve the resulting unconstrained problem which helps reduce the complexity significantly, compared to conventional methods. Simulation results show the fast convergence of the proposed algorithm and substantial performance improvements over existing approaches.

Index Terms—Cognitive radio, interference, jamming noise, physical layer security, zero-forcing, transmit beamforming.

I. INTRODUCTION

Physical layer (PHY)-security has been considered as a promising technique to prevent eavesdropping without upper layer data encryption. The key aim of PHY-security is to guarantee a positive secrecy rate of the legitimate user by exploiting random characteristics of the wireless channel [2], [3]. To make PHY-security viable, the quality of the legitimate channel is required to be better than the eavesdroppers' [4]. However, this assumption may not be always feasible

in practice. Traditionally, one possible way to improve the secrecy rate is to use multiple antennas at the transmitter for designing beamforming vectors [5]–[7]. This technique allows the concentration of the transmit signal over the intended user's direction while reducing power leakage to the eavesdroppers. Furthermore, the transmitter tries to debilitate the desired signal at the eavesdroppers.

Recently, another promising technique has been proposed to embed jamming noise (JN), also known as artificial noise, at the transmitter, which is transmitted simultaneously with its own information signal to degrade the channel of the eavesdropper. Notably, [8] first introduced the transmitter that spends some of the available power to produce the JN for the purpose of degrading the channel to eavesdroppers. The use of JN for beamforming design can be divided into two major categories, depending on how much the transmitter knows the channel state information (CSI) of the eavesdropper, e.g., with passive and active eavesdroppers. In [9], a passive eavesdropper case was considered where the JN is designed to null out the interference to the legitimate user. Inspired by the work in [9], the authors in [10]–[13] analyzed and optimized the secrecy performance of the systems. The common design of those works is to force JN beamforming into the null space of the legitimate channel, whereas the beamforming for transmitting information is fixed to maximum-ratio-transmission (MRT). For the case of active eavesdroppers, the secrecy rate maximization problem for multiple-input single-output (MISO) channel overheard by multiple eavesdroppers was considered in [14], [15], where the rank relaxation is proved to be tight by applying a semidefinite program (SDP) relaxation. In addition, the imperfect knowledge of the eavesdropper's CSI at the legitimate transmitter was presented as an emerging subject in different works, e.g., related to the outage robust design [16] and the worst-case robust transmit design [17], [18].

Moreover, similarly to other wireless networks, cognitive radio networks (CRNs) are particularly faced with security threats due to the openness of wireless transmission media. Although security is a critical issue, PHY-security of CRNs has not been well studied until recently. Several efforts of PHY-security in CRNs have been reported in [5], [19]–[22]. In [19], by considering a secondary transmitter equipped with multiple antennas, the authors designed a beamforming vector to maximize the secrecy capacity of the secondary system, as well as to help the primary system improve the secrecy

V.-D. Nguyen and O.-S. Shin are with the School of Electronic Engineering, Soongsil University, Seoul 06978, Korea (e-mail: {nguyenvandinh, osshin}@ssu.ac.kr).

T. Q. Duong is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, United Kingdom (e-mail: trung.q.duong@qub.ac.uk).

O. A. Dobre is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL, Canada (e-mail: odobre@mun.ca).

Part of this work was presented at the 2016 IEEE International Conference on Communications (ICC) [1].

This work was supported in part by Basic Research Laboratories (BRL) through NRF grant funded by the MSIP (No. 2015056354). This work was supported in part by the U.K. Royal Academy of Engineering Fellowship under Grant RF1415\14\22 and by the Newton Institutional Link under Grant ID 172719890.

capacity. A cooperative communication between the secondary and primary systems was studied in [20] to improve the secrecy capacity of the primary system while satisfying the quality of service (QoS) of the secondary system. The optimal robust design problem for secure MISO CRNs, where the secondary transmitter does not have perfect CSI of all the channels but only knows the uncertainty regions containing the actual channels, was addressed in [21]. Furthermore, the primary transmitter with a jamming beamforming that avoids interference to the primary receiver was investigated in [22], where a closed-form expression of the achievable rate was derived. Most of the prior research on beamforming design for CRNs assumed either a single primary receiver or a single eavesdropper. In addition, the CSI of the eavesdropper's channel was assumed to be perfectly known at the transmitter. However, those assumptions are not practical, particularly when the eavesdroppers are passive devices.

To overcome these limitations of the previous works, in this paper we consider the case where multiple eavesdroppers wiretap the confidential messages from the secondary system in the presence of multiple primary users. Different from the works in [5], [19]–[21], we investigate the scenario where the CSI of the eavesdroppers' channels is partially known [9], [22], in the statistical sense. It is important to note that the results in [9], [22] cannot be directly applied to our scenario since the transmit and JN beamformings are fixed and mainly focus on power allocation. In addition, the beamforming design of the works in [9], [22] may not be efficient for an underlay CRN, where the interferences to the primary users caused by the secondary system cannot exceed a certain threshold. For the present problem, we consider the joint optimization of information and JN beamforming to maximize the secrecy rate subject to per-antenna power constraints (PAPCs) at the secondary transmitter and the interference power constraint at the primary users. The problem of interest differs from the ones studied previously, which often assumed a sum power constraint. In many existing approaches, the beamforming designs for information and JN are not simultaneously optimized. In this paper we adopt the zero-forcing method at the secondary transmitter to eliminate the interference caused by JN to the secondary user, for its simplicity and effectiveness. To facilitate the maximization of the secrecy rate, we introduce an auxiliary variable to control the level of mutual information from the secondary transmitter to eavesdroppers. For an optimal solution, the beamforming design problem can be cast as an SDP by showing that the rank relaxation is tight as in [14], [15], [19], which can be solved by generic conic solvers such as SDPT3 [30] and SeDuMi [31]. However, we do not follow such an approach for the following two reasons. First, it provides few useful insights into the structure of the optimal beamforming design. Second, its computational complexity is generally very high when the problem size becomes large for a real-time implementation, and it has no explicit form, thus hindering any further theoretical analysis.

In this paper, we propose an efficient algorithm for the optimization problem by first transforming the constraint resulting from the statistical information of the passive eavesdroppers

into a linear matrix inequality and convex constraint, and prove the optimality of the rank relaxation. By extending a duality result, we convert the considered problem into a minimax program with an equality constraint, for which a barrier method is derived to find a saddle point. The proposed algorithm is an iterative Newton method which exploits the special features of the design problem. In particular, in each iteration to find the Newton step, we apply a block elimination method to derive a system of generalized Sylvester equations; this requires a reduced complexity compared to a generic method based on solving a system of linear equations. To find a global optimal solution, we design an optimization problem via the dual problem associated with the algorithm of quasi-Newton optimization, referred to as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. In fact, the BFGS method is an approximation of the Newton method for solving non-linear problems, and often exhibits a superior convergence rate.

The rest of this paper is organized as follows. The system model and the formulation of the secrecy rate maximization problem are described in Section II. In Section III, we derive the optimization problem for the beamforming design. In Section IV we present the proposed algorithms to solve this problem. Numerical results are provided in Section V, and Section VI concludes the paper.

II. SYSTEM MODEL AND NOTATION

In this section, after introducing the notation used in this paper, we present a CRN model for secure communication with concurrent wireless information and jamming noise transfer.

A. Notation

Bold lower and upper case letters represent vectors and matrices, respectively. \mathbf{X}^H , \mathbf{X}^T , $\text{tr}(\mathbf{X})$, and $\text{rank}(\mathbf{X})$ are the Hermitian transpose, normal transpose, trace, and rank of a matrix \mathbf{X} , respectively. $\|\cdot\|$ and $|\cdot|$ denote the Euclidean norm of a matrix or vector and the absolute value of a complex scalar, respectively. \mathbf{I}_N represents an $N \times N$ identity matrix. $[\mathbf{x}]_i$ is the i -th entry of vector \mathbf{x} . $[\mathbf{X}]_{i,j}$ is the entry at the i -th row and j -th column of \mathbf{X} . \mathbf{e}_n is the n -th unit vector, i.e., $[\mathbf{e}_n]_n = 1$ and $[\mathbf{e}_n]_i = 0$, $\forall i \neq n$. $\text{diag}(\mathbf{x})$, where \mathbf{x} is a vector, denotes a diagonal matrix with diagonal elements of \mathbf{x} . $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\eta}, \mathbf{Z})$ indicates that \mathbf{x} is a random vector following a complex circularly symmetric Gaussian distribution with mean vector $\boldsymbol{\eta}$ and covariance matrix \mathbf{Z} . The notation $\mathbf{X} \succeq \mathbf{0}$ represents a positive semidefinite matrix \mathbf{X} . \mathbb{H}^N denotes the set of $N \times N$ complex Hermitian matrices. $\lambda_{\max}(\mathbf{X})$ and $\lambda_j(\mathbf{X})$ denote the maximum eigenvalue and the j -th eigenvalue of a Hermitian matrix \mathbf{X} , respectively. $\nabla_{\mathbf{x}} f(\mathbf{x})$ represents the gradient of $f(\cdot)$ with respect to vector \mathbf{x} . The acronyms inf and sup denote the terms of infimum and supremum, respectively.

B. Signal Model

We consider the PHY-security of a CRN consisting of one secondary transmitter (ST), one secondary receiver (SU), M primary users (PUs), and K eavesdroppers (Eves), as

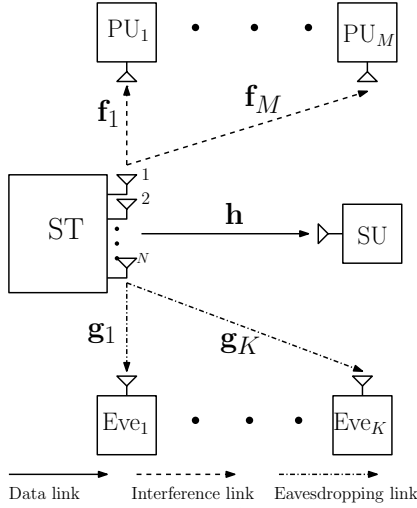


Fig. 1. A CRN model with multiple eavesdroppers.

illustrated in Fig. 1. The ST is equipped with N antennas, while the other nodes are equipped with a single antenna. In the secondary system, Eves intend to wiretap and decode confidential messages from the ST. All the channels are assumed to remain constant during a transmission block and change independently from one block to another.

The main objective of the design is to maximize the secrecy rate of the SU while satisfying a given interference power constraint I_m at the m -th PU, for $m = 1, 2, \dots, M$. We aim to design two beamforming vectors \mathbf{w} and \mathbf{u} at the ST, corresponding to the data and the jamming noise as

$$\mathbf{x}_s = \mathbf{w}s_c + \mathbf{u} \quad (1)$$

where $s_c \in \mathbb{C}$ is the confidential message that the ST transmits to SU, with $\mathbb{E}\{|s_c|^2\} = 1$, which is weighted with the beamforming vector $\mathbf{w} \in \mathbb{C}^{N \times 1}$, and \mathbf{u} is the jamming noise vector whose elements are zero-mean complex Gaussian random variables with covariance matrix \mathbf{U} , i.e., $\mathbf{u} \sim \mathcal{CN}(\mathbf{0}, \mathbf{U})$, where $\mathbf{U} \in \mathbb{H}^N$ and $\mathbf{U} \succeq \mathbf{0}$.

The signal-to-interference-plus-noise ratio (SINR) at the SU and at the k -th Eve are, respectively, given by

$$\Gamma_s = \frac{|\mathbf{h}^H \mathbf{w}|^2}{\mathbf{h}^H \mathbf{U} \mathbf{h} + 1} \text{ and } \Gamma_k = \frac{|\mathbf{g}_k^H \mathbf{w}|^2}{\mathbf{g}_k^H \mathbf{U} \mathbf{g}_k + 1}, \forall k \in \mathcal{K} \quad (2)$$

where $\mathbf{h} \in \mathbb{C}^{N \times 1}$ and $\mathbf{g}_k \in \mathbb{C}^{N \times 1}$ are the baseband equivalent channels of the links from the ST to SU and to the k -th Eve, respectively. \mathcal{K} is defined as $\mathcal{K} \triangleq \{1, 2, \dots, K\}$. Without loss of generality, the background thermal noise at each receiver is assumed to be a zero-mean and unit variance complex Gaussian random variable.

C. Problem Formulation

The secrecy rate of the SU, R_s , is defined as [24]

$$R_s = \max \left(\log(1 + \Gamma_s) - \max_{k \in \mathcal{K}} \log(1 + \Gamma_k), 0 \right). \quad (3)$$

If R_s is kept larger than zero while ensuring that the received interference at the m -th PU is below the predetermined threshold I_m , then the signal transmitted from the ST to the SU is “undecodable,” as indicated in [25].

In this paper, the beamformer for the jamming noise is designed to null out the interference to SU, such that

$$\Phi^H \mathbf{h} = \mathbf{0} \quad (4)$$

where we choose $\mathbf{U} = \Phi \Phi^H$ with $\Phi \in \mathbb{C}^{N \times (N-1)}$ [9], [22]. Thus, the SINR of SU and the k -th Eve can be rewritten as

$$\Gamma_s = |\mathbf{h}^H \mathbf{w}|^2 \text{ and } \Gamma_k = \frac{|\mathbf{g}_k^H \mathbf{w}|^2}{\|\mathbf{g}_k^H \Phi\|^2 + 1}, \forall k \in \mathcal{K}. \quad (5)$$

Remark 1: The assumption of null space jamming noise in (4) is reasonable, since the use of jamming noise is intended to degrade the channel of eavesdroppers, but not the legitimate user [9], [22]. Importantly, the use of the null space jamming noise constraint may simplify the design and analysis [16], [17].

The optimization problem can be formulated as

$$\mathbf{P1} : \max_{\mathbf{w}, \Phi} \left\{ \log(1 + \Gamma_s) - \max_{k \in \mathcal{K}} \log(1 + \Gamma_k) \right\} \quad (6a)$$

$$\text{s. t. } \Phi^H \mathbf{h} = \mathbf{0} \quad (6b)$$

$$[\mathbf{w} \mathbf{w}^H]_{n,n} + [\Phi \Phi^H]_{n,n} \leq P_n, \forall n \in \mathcal{N} \quad (6c)$$

$$|\mathbf{f}_m^H \mathbf{w}|^2 + \|\mathbf{f}_m^H \Phi\|^2 \leq I_m, \forall m \in \mathcal{M} \quad (6d)$$

where $\mathcal{N} \triangleq \{1, 2, \dots, N\}$ and $\mathcal{M} \triangleq \{1, 2, \dots, M\}$. $\mathbf{f}_m \in \mathbb{C}^{N \times 1}$ is the baseband equivalent channel of the link from the ST to the m -th PU. The constraint in (6c) represents the power constraint for the n -th antenna at the ST. We note that each antenna is often equipped with its own power amplifier (PA). Thus, one may need to limit the per-antenna peak power to operate within the linear region of the PA [26]. The PAPCs in (6c) are different from the sum power constraint (SPC) considered in [15], [20]; however, the proposed beamforming scheme in this paper can also be applied to the SPC with slight modifications. The constraint in (6d) is to protect the primary system, so that the interference power at the m -th PU due to the ST is less than a given interference threshold I_m , $\forall m \in \mathcal{M}$.

By introducing an auxiliary variable Γ_{tol} and in the spirit of [27], $\mathbf{P1}$ has the same optimal solutions as the following new problem

$$\mathbf{P2} : \max_{\mathbf{w}, \Phi, \Gamma_{\text{tol}} > 0} \log(1 + |\mathbf{h}^H \mathbf{w}|^2) - \log(1 + \Gamma_{\text{tol}}) \quad (7a)$$

$$\text{s. t. } \max_{k \in \mathcal{K}} \frac{|\mathbf{g}_k^H \mathbf{w}|^2}{\|\mathbf{g}_k^H \Phi\|^2 + 1} \leq \Gamma_{\text{tol}} \quad (7b)$$

$$(6b), (6c), (6d) \quad (7c)$$

where $\Gamma_{\text{tol}} > 0$ is the maximum allowable SINR for Eves to wiretap the confidential messages from the ST. Intuitively, we have an equivalent problem with less difficulty by adjusting Γ_{tol} .

To further simplify $\mathbf{P2}$, let $\bar{\mathbf{V}} \in \mathbb{C}^{N \times (N-1)}$ be the null space of \mathbf{h}^H . Then we can write $\Phi = \bar{\mathbf{V}} \bar{\Phi}$, where $\bar{\Phi} \in$

$\mathbb{C}^{(N-1) \times (N-1)}$ is the solution to the following problem

$$\mathbf{P3} : \max_{\mathbf{w}, \bar{\Phi}, \Gamma_{\text{tol}} > 0} \log(1 + |\mathbf{h}^H \mathbf{w}|^2) - \log(1 + \Gamma_{\text{tol}}) \quad (8a)$$

$$\text{s. t. } \max_{k \in \mathcal{K}} \frac{|\mathbf{g}_k^H \mathbf{w}|^2}{\mathbf{g}_k^H \bar{\mathbf{V}} \bar{\Phi} \bar{\Phi}^H \bar{\mathbf{V}}^H \mathbf{g}_k + 1} \leq \Gamma_{\text{tol}} \quad (8b)$$

$$[\mathbf{w} \mathbf{w}^H]_{n,n} + [\bar{\mathbf{V}} \bar{\Phi} \bar{\Phi}^H \bar{\mathbf{V}}^H]_{n,n} \leq P_n, \forall n \in \mathcal{N} \quad (8c)$$

$$|\mathbf{f}_m^H \mathbf{w}|^2 + \|\mathbf{f}_m^H \bar{\mathbf{V}} \bar{\Phi}\|^2 \leq I_m, \forall m \in \mathcal{M}. \quad (8d)$$

D. Channel State Information

We consider the case that the CSI of \mathbf{h} and \mathbf{f}_m , $\forall m$, is perfectly known at the ST [5], [19], [20], where the SU and PUs are active users. Explicitly, the ST sends pilot signals at the beginning of each scheduling slot to the SU and PUs. The channel vectors are estimated at the SU and PUs and then fed back to the ST using a dedicated control channel. In this paper, we assume that Eves passively wiretap the confidential messages transmitted from the ST to SU without causing any interference to the SU and PUs. For the passive Eves, the entries of \mathbf{g}_k , $\forall k$, are modeled as independent and identically distributed (i.i.d.) Rayleigh fading channels, where the instantaneous information of these wiretap channels are not available at ST. These assumptions about the passive Eves are commonly used in the literature [9], [22], [23]. Based on the above settings, **P3** can be rewritten as

$$\mathbf{P4} : \max_{\mathbf{w}, \bar{\Phi}, \Gamma_{\text{tol}} > 0} \log(1 + |\mathbf{h}^H \mathbf{w}|^2) - \log(1 + \Gamma_{\text{tol}}) \quad (9a)$$

$$\text{s. t. } \Pr \left(\max_{k \in \mathcal{K}} \frac{\mathbf{w}^H \mathbf{G}_k \mathbf{w}}{\text{tr}(\mathbf{G}_k \bar{\mathbf{V}} \bar{\Phi} \bar{\Phi}^H \bar{\mathbf{V}}^H) + 1} \leq \Gamma_{\text{tol}} \right) \geq \kappa \quad (9b)$$

$$[\mathbf{w} \mathbf{w}^H]_{n,n} + [\bar{\mathbf{V}} \bar{\Phi} \bar{\Phi}^H \bar{\mathbf{V}}^H]_{n,n} \leq P_n, \forall n \in \mathcal{N} \quad (9c)$$

$$|\mathbf{f}_m^H \mathbf{w}|^2 + \|\mathbf{f}_m^H \bar{\mathbf{V}} \bar{\Phi}\|^2 \leq I_m, \forall m \in \mathcal{M}. \quad (9d)$$

where $\mathbf{G}_k \triangleq \mathbf{g}_k \mathbf{g}_k^H$, and κ is a parameter for providing secure communication. In particular, the maximum received SINR at all Eves is required to be less than a given value Γ_{tol} with at least probability κ [23].

III. OPTIMIZATION PROBLEM DESIGN

The global optimal Γ_{tol} of **P4** can be found from one dimensional search. Therefore, our main task in this section is to derive a convex optimization approach to **P4** with respect to \mathbf{w} and $\bar{\Phi}$ for a fixed value of Γ_{tol} . A method to find the optimal solution of Γ_{tol} will be presented in the next section.

We note that **P4** is not a convex program. A standard way to solve (9) for a fixed Γ_{tol} is to consider the following problem

$$\mathbf{P5} : \max_{\mathbf{W} \succeq \mathbf{0}, \bar{\mathbf{U}} \succeq \mathbf{0}} \log(1 + \mathbf{h}^H \mathbf{W} \mathbf{h}) \quad (10a)$$

$$\text{s. t. } \Pr \left(\max_{k \in \mathcal{K}} \frac{\text{tr}(\mathbf{G}_k \mathbf{W})}{\text{tr}(\mathbf{G}_k \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H) + 1} \leq \Gamma_{\text{tol}} \right) \geq \kappa \quad (10b)$$

$$[\mathbf{W}]_{n,n} + [\bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H]_{n,n} \leq P_n, \forall n \in \mathcal{N} \quad (10c)$$

$$\text{tr}(\mathbf{F}_m \mathbf{W}) + \text{tr}(\bar{\mathbf{F}}_m \bar{\mathbf{U}}) \leq I_m, \forall m \in \mathcal{M} \quad (10d)$$

$$\text{rank}(\mathbf{W}) = 1 \quad (10e)$$

where $\mathbf{W} \triangleq \mathbf{w} \mathbf{w}^H$, $\bar{\mathbf{U}} \triangleq \bar{\Phi} \bar{\Phi}^H$, $\mathbf{F}_m \triangleq \mathbf{f}_m \mathbf{f}_m^H$, $\bar{\mathbf{F}}_m \triangleq \bar{\mathbf{f}}_m \bar{\mathbf{f}}_m^H$, and $\bar{\mathbf{f}}_m \triangleq \bar{\mathbf{V}}^H \mathbf{f}_m$. In addition, we have dropped $\log(1 + \Gamma_{\text{tol}})$

from the objective function in (10) to have a simpler problem without affecting optimality. We remark that the constraint $\text{rank}(\mathbf{W}) = 1$ must be satisfied to transmit the confidential message $s_c \in \mathbb{C}$. **P5** is still a nonconvex program due to the nonconvex constraints in (10b) and (10e).

To make **P5** a tractable problem, we first transform the constraint in (10b) into a linear matrix inequality and convex constraint according to the following lemma.

Lemma 1: The constraint in (10b) can be transformed as

$$\mathbf{W} - \Gamma_{\text{tol}} \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H \preceq \mathbf{I} \xi \quad (11)$$

where $\xi = \Phi_N^{-1}(1 - \kappa^{1/K}) \Gamma_{\text{tol}}$, with $\Phi_N^{-1}(\cdot)$ being the inverse cumulative distribution function of an inverse central chi-square random variable with $2N$ degrees of freedom.

Proof: See Appendix A. ■

Remark 2: Note that the implication in (11) can be applied to any continuous channel distribution by replacing $\Phi_N^{-1}(\cdot)$ with the corresponding one. Consequently, the proposed solution introduced in this paper also applies to other eavesdroppers' channel distributions with slight modification of the optimization problem.

By replacing (10b) with (11), we obtain the following new problem

$$\mathbf{P6} : \max_{\mathbf{W}, \bar{\mathbf{U}}} \log(1 + \mathbf{h}^H \mathbf{W} \mathbf{h}) \quad (12a)$$

$$\text{s. t. } \mathbf{W} - \Gamma_{\text{tol}} \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H \preceq \mathbf{I} \xi \quad (12b)$$

$$[\mathbf{W}]_{n,n} + [\bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H]_{n,n} \leq P_n, \forall n \in \mathcal{N} \quad (12c)$$

$$\text{tr}(\mathbf{F}_m \mathbf{W}) + \text{tr}(\bar{\mathbf{F}}_m \bar{\mathbf{U}}) \leq I_m, \forall m \in \mathcal{M} \quad (12d)$$

$$\mathbf{W} \succeq \mathbf{0}, \bar{\mathbf{U}} \succeq \mathbf{0}, \text{rank}(\mathbf{W}) = 1. \quad (12e)$$

We note that the feasible solutions of (12) also satisfy (10) but not vice versa due to the inequality in (59). In other words, (11) is a relaxation of (10b) which yields a large feasible solution set for **P5**. Although (12) is a nonconvex program, it can be efficiently solved with some numerical solvers by dropping the rank constraint in (12e); then, the considered problem **P6** becomes a so-called rank relaxed problem. Importantly, we prove that the rank relaxation is tight in the following lemma.

Lemma 2: The optimal solution \mathbf{W}^* to (12) satisfies $\text{rank}(\mathbf{W}^*) = 1$.

Proof: See Appendix B. ■

Remark 3: We note that the optimization problem with a fixed Γ_{tol} in (12) is also applicable for the SINR-based design in [19], [23]. More challengingly, the goal of this paper is to provide a secrecy rate maximization for the secondary system, rather than a certain quality-of-service.

The problem **P6** now is a convex SDP, which can be solved by numerical solvers such as SDPT3 [30] and SeDuMi [31]. However, solving (12) directly may not be efficient since \mathbf{W} is treated as a Hermitian matrix of N^2 real variables, which is of relatively high complexity especially when N becomes large. As mentioned earlier, such a method is not considered in this paper. We now present the following theorem.

Theorem 1: Consider the following minimax problem

$$\begin{aligned} \mathbf{P7} : \quad & \min_{\psi \geq 0, \mathbf{D} \succeq \mathbf{0}} \max_{w \geq 0, \dot{\mathbf{U}} \succeq \mathbf{0}} \log \frac{|\Sigma + \mathbf{h}w\mathbf{h}^H|}{|\Sigma|} \\ & \text{s. t. } w + \text{tr}(\Omega \dot{\mathbf{U}}) \leq P \\ & \xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi \leq P \end{aligned} \quad (13)$$

where $\Sigma = \mathbf{D} + \text{diag}(\lambda) + \sum_{m=1}^M \mu_m \mathbf{F}_m$, $\Omega = \bar{\mathbf{V}}^H \left(-\Gamma_{\text{tol}} \mathbf{D} + \text{diag}(\lambda) + \sum_{m=1}^M \mu_m \mathbf{F}_m \right) \bar{\mathbf{V}}$, and $\psi = [\lambda^T \ \mu^T]^T$. Then, the optimal solution \mathbf{W}^* of the relaxed problem (12) can be obtained from that of (13) as

$$\mathbf{W} = \frac{\Sigma^{-1} \mathbf{h}w\mathbf{h}^H \Sigma^{-1}}{\mathbf{h}^H \Sigma^{-1} \mathbf{h}}. \quad (14)$$

Proof: See Appendix C. ■

We remark that the problem in (13) is equivalent to (12), and the inequality constraints in (13) hold with equality at optimality. We now provide a sketch of the proof to verify this point. Suppose $w + \text{tr}(\Omega \dot{\mathbf{U}}) < P$ for a fixed (ψ, \mathbf{D}) , there possibly exists an arbitrarily small value $\epsilon > 0$ that satisfies $w + \epsilon + \text{tr}(\Omega \dot{\mathbf{U}}) < P$. Replacing w by $w + \epsilon/2$ in the objective function in (13) yields a larger objective value, which leads to contradiction since w is optimal. A similar observation can be made in the constraint $\xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi \leq P$ for a fixed $(w, \dot{\mathbf{U}})$. Thus, (13) can be efficiently solved by the interior-point methods in [33] since the equality constraints are easy to handle in general. Interestingly, solving (13) for the optimal solution of $w \in \mathbb{C}$ requires much lower complexity, when compared with solving $\mathbf{W} \in \mathbb{C}^{N \times N}$. We propose a numerical algorithm to solve (13) based on interior-point methods in the next section.

Remark 4: As mentioned in Section II-C, we now show how **P7** can be modified to include a SPC which is written as $\text{tr}(\mathbf{W} + \bar{\mathbf{V}} \dot{\mathbf{U}} \bar{\mathbf{V}}^H) \leq P$, where P is the total transmit power at the ST. By following the same steps as presented in Appendix C, we can arrive at a similar minimax problem as shown in **P7**, where Σ , Ω , and ψ are changed to $\Sigma = \mathbf{D} + \lambda \mathbf{I} + \sum_{m=1}^M \mu_m \mathbf{F}_m$, $\Omega = \bar{\mathbf{V}}^H \left(-\Gamma_{\text{tol}} \mathbf{D} + \lambda \mathbf{I} + \sum_{m=1}^M \mu_m \mathbf{F}_m \right) \bar{\mathbf{V}}$, and $\psi = [\lambda \ \mu^T]^T$, respectively, and $\lambda \in \mathbb{C}$ is the dual variable associated with the SPC. Consequently, Algorithm 1 is rather general in the sense that it can be applied to handle this case as well. In addition, if we assume that the total transmit power at the ST is the sum of all PAPCs, i.e., mathematically $P = \sum_{n=1}^N P_n$, then it follows that $R_s^{\text{SPC}} \geq R_s^{\text{PAPCs}}$, where R_s^{SPC} and R_s^{PAPCs} are the optimal values of (3) corresponding to the SPC and PAPCs, respectively.

IV. PROPOSED ALGORITHM

A. Proposed Algorithm to Solve (13) with a fixed Γ_{tol}

For simplicity, we define $\bar{\mathbf{U}} = \Omega^{-1/2} \dot{\mathbf{U}} \Omega^{-1/2}$ since Ω is invertible and then the relationship between \mathbf{U} and $\bar{\mathbf{U}}$ is given by

$$\mathbf{U} = \bar{\mathbf{V}} \Omega^{-1/2} \dot{\mathbf{U}} \Omega^{-1/2} \bar{\mathbf{V}}^H. \quad (15)$$

We now present a computationally efficient algorithm to solve (13) to find the optimal solution for w and $\bar{\mathbf{U}}$. Toward this end, we use the relationships in (14) and (15) to derive the

optimal beamformers of \mathbf{W} and \mathbf{U} . As mentioned earlier, we consider the following problem rather than problem (13) without affecting its optimality.

$$\begin{aligned} \min_{\psi \geq 0, \mathbf{D} \succeq \mathbf{0}} \max_{w \geq 0, \dot{\mathbf{U}} \succeq \mathbf{0}} \quad & \log \frac{|\Sigma + \mathbf{h}w\mathbf{h}^H|}{|\Sigma|} \\ \text{s. t. } \quad & w + \text{tr}(\dot{\mathbf{U}}) = P \\ & \xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi = P. \end{aligned} \quad (16)$$

The proposed method to solve (16) is based on a barrier method to find a saddle point [33]. Accordingly, we consider the modified objective function given by

$$\begin{aligned} f(t, \psi, \mathbf{D}, w, \dot{\mathbf{U}}) = & \log \frac{|\Sigma + \mathbf{h}w\mathbf{h}^H|}{|\Sigma|} - \frac{1}{t} \sum_{i=1}^{N+M} \log(\psi_i) \\ & - \frac{1}{t} \log |\mathbf{D}| + \frac{1}{t} \log(w) + \frac{1}{t} \log |\dot{\mathbf{U}}| \end{aligned} \quad (17)$$

where $\log |\dot{\mathbf{U}}|$, $\log |\mathbf{D}|$, $\log(\psi_i)$, and $\log(w)$ are the logarithmic barrier functions to account for the constraints $\dot{\mathbf{U}} \succeq \mathbf{0}$, $\mathbf{D} \succeq \mathbf{0}$, $\psi_i \geq 0$, and $w \geq 0$, respectively, and $t > 0$ is a parameter that controls the logarithm barrier terms. We should remark that the modified objective function (17) is concave in $(w, \dot{\mathbf{U}})$ for a fixed value of (ψ, \mathbf{D}) , and convex in (ψ, \mathbf{D}) for a fixed value of $(w, \dot{\mathbf{U}})$. For a given value of t , the barrier method requires to solve a standard equality constrained maximization problem

$$\min_{\psi \geq 0, \mathbf{D} \succeq \mathbf{0}} \max_{w \geq 0, \dot{\mathbf{U}} \succeq \mathbf{0}} f(t, \psi, \mathbf{D}, w, \dot{\mathbf{U}}) \quad (18a)$$

$$\text{s. t. } w + \text{tr}(\dot{\mathbf{U}}) = P \quad (18b)$$

$$\xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi = P. \quad (18c)$$

The main idea of the interior-point method used to solve (18) is summarized as follows: we find the optimal solution $(\psi, \mathbf{D}, w, \dot{\mathbf{U}})$ for a fixed value t (which is referred to as a centering step), and increase t until the dual gap of the equality constrained maximization problem (18) satisfies a given accuracy. To do this, we start with the Karush-Kuhn-Tucker (KKT) conditions for (18), which are given by

$$\mathbf{h}^H (\Sigma + \mathbf{h}w\mathbf{h}^H)^{-1} \mathbf{h} + \frac{1}{t} w^{-1} - \eta_1 = 0. \quad (19a)$$

$$\frac{1}{t} \dot{\mathbf{U}}^{-1} - \eta_1 \mathbf{I} = \mathbf{0}. \quad (19b)$$

$$w + \text{tr}(\dot{\mathbf{U}}) = P. \quad (19c)$$

$$\begin{aligned} \mathbf{e}_n^H [(\Sigma + \mathbf{h}w\mathbf{h}^H)^{-1} - \Sigma^{-1}] \mathbf{e}_n \\ - \frac{\lambda_n^{-1}}{t} + \eta_2 P_n = 0, \forall n. \end{aligned} \quad (19d)$$

$$\begin{aligned} \mathbf{f}_m^H [(\Sigma + \mathbf{h}w\mathbf{h}^H)^{-1} - \Sigma^{-1}] \mathbf{f}_m \\ - \frac{\mu_m^{-1}}{t} + \eta_2 I_m = 0, \forall m. \end{aligned} \quad (19e)$$

$$(\Sigma + \mathbf{h}w\mathbf{h}^H)^{-1} - \Sigma^{-1} - \frac{1}{t} \mathbf{D}^{-1} + \eta_2 \xi \mathbf{I} = \mathbf{0}. \quad (19f)$$

$$\xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi = P. \quad (19g)$$

where η_1 and η_2 are the dual variables corresponding to the constraints in (18b) and (18c), respectively. The results in (19a), (19b), (19d), (19e), and (19f) are obtained by utilizing

the gradient of the Lagrangian function with respect to w , $\dot{\mathbf{U}}$, λ , η , and \mathbf{D} , respectively.

Toward this end, we use the infeasible start Newton method to find the optimal solution of (19). The aim of the Newton method is to calculate a Newton step in each iteration. Specifically, we replace w by $w + \Delta w$, $\dot{\mathbf{U}}$ by $\dot{\mathbf{U}} + \Delta \dot{\mathbf{U}}$, λ by $\lambda + \Delta \lambda$, μ by $\mu + \Delta \mu$, \mathbf{D} by $\mathbf{D} + \Delta \mathbf{D}$, and η_i by $\eta_i + \Delta \eta_i$, $i \in \{1, 2\}$, in (19a) to have

$$t\mathbf{h}^H(\Sigma + \mathbf{h}\mathbf{w}\mathbf{h}^H + \Delta\Sigma + \mathbf{h}\Delta w\mathbf{h}^H)^{-1}\mathbf{h} \\ + (w + \Delta w)^{-1} - t(\eta_1 + \Delta\eta_1) = 0 \quad (20)$$

where $\Delta\Sigma = \Delta\mathbf{D} + \text{diag}(\Delta\lambda) + \sum_{m=1}^M \Delta\mu_m \mathbf{F}_m$. By applying the identity $(\mathbf{A} + \mathbf{B})^{-1} \approx \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$ for small \mathbf{B} ,¹ (20) can be approximated as

$$t\mathbf{h}^H\Xi\Delta\Sigma\Xi\mathbf{h}w^2 + t\mathbf{h}^H\Xi\mathbf{h}\Delta w\mathbf{h}^H\Xi\mathbf{h}w^2 + \Delta w \\ + tw^2\Delta\eta_1 = tw\mathbf{h}^H\Xi\mathbf{h}w + w - tw^2\eta_1 \quad (21)$$

where $\Xi \triangleq (\Sigma + \mathbf{h}\mathbf{w}\mathbf{h}^H)^{-1}$. Note that

$$\mathbf{h}^H\Xi\Delta\Sigma\Xi\mathbf{h} = \mathbf{h}^H\Xi\Delta\mathbf{D}\Xi\mathbf{h} + \sum_{n=1}^N \Delta\lambda_n \mathbf{h}^H\Xi\mathbf{e}_n\mathbf{e}_n^H\Xi\mathbf{h} \\ + \sum_{m=1}^M \Delta\mu_m \mathbf{h}^H\Xi\mathbf{F}_m\Xi\mathbf{h}. \quad (22)$$

Substituting (22) into (21), we arrive at

$$t\mathbf{h}^H\Xi\Delta\mathbf{D}\Xi\mathbf{h}w^2 + t \sum_{n=1}^N \Delta\lambda_n \mathbf{h}^H\Xi\mathbf{e}_n\mathbf{e}_n^H\Xi\mathbf{h}w^2 \\ + t \sum_{m=1}^M \Delta\mu_m \mathbf{h}^H\Xi\mathbf{F}_m\Xi\mathbf{h}w^2 + t\mathbf{h}^H\Xi\mathbf{h}\Delta w\mathbf{h}^H\Xi\mathbf{h}w^2 \\ + \Delta w + tw^2\Delta\eta_1 = tw\mathbf{h}^H\Xi\mathbf{h}w + w - tw^2\eta_1. \quad (23)$$

Next, from (19d), (19e), and (19f), we have

$$t\mathbf{e}_n^H[(\Sigma + \mathbf{h}\mathbf{w}\mathbf{h}^H + \Delta\Sigma + \mathbf{h}\Delta w\mathbf{h}^H)^{-1} - (\Sigma + \Delta\Sigma)^{-1}]\mathbf{e}_n \\ - (\lambda_n + \Delta\lambda_n)^{-1} + t(\eta_2 + \Delta\eta_2)P_n = 0, \forall n. \quad (24)$$

$$t\mathbf{f}_m^H[(\Sigma + \mathbf{h}\mathbf{w}\mathbf{h}^H + \Delta\Sigma + \mathbf{h}\Delta w\mathbf{h}^H)^{-1} - (\Sigma + \Delta\Sigma)^{-1}]\mathbf{f}_m \\ - (\mu_m + \Delta\mu_m)^{-1} + t(\eta_2 + \Delta\eta_2)I_m = 0, \forall m. \quad (25)$$

$$t(\Sigma + \mathbf{h}\mathbf{w}\mathbf{h}^H + \Delta\Sigma + \mathbf{h}\Delta w\mathbf{h}^H)^{-1} - t(\Sigma + \Delta\Sigma)^{-1} \\ - (\mathbf{D} + \Delta\mathbf{D})^{-1} + t(\eta_2 + \Delta\eta_2)\xi\mathbf{I} = \mathbf{0}. \quad (26)$$

Following the same steps from (21) to (23), (24) can be approximated as

$$t\mathbf{e}_n^H\Xi\Delta\mathbf{D}\Xi\mathbf{e}_n - t\mathbf{e}_n^H\Sigma^{-1}\Delta\mathbf{D}\Sigma^{-1}\mathbf{e}_n + t \sum_{i=1}^N \Delta\lambda_i\varphi_{n,i} \\ + t \sum_{m=1}^M \Delta\mu_m\varphi_{n,m} + t\Delta w\phi_n - \Delta\lambda_n\lambda_n^{-2} - t\Delta\eta_2P_n \\ = t\varphi_n - \lambda_n^{-1} + t\eta_2P_n, \forall n \quad (27)$$

¹The approximation $(\mathbf{A} + \mathbf{B})^{-1} \approx \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$ is proper for small entries of matrix \mathbf{B} and relatively rough for large entries of matrix \mathbf{B} [34, Chapter 3]. In particular, the residual error, as shown in Fig. 2, is high for the first iterations since the approximation is not reliable. However, when Algorithm 1 reaches the optimal solution due to updating \mathbf{B} in each iteration, the residual error will be small and thus the approximation becomes very accurate.

where $\varphi_{n,i} \triangleq |\mathbf{e}_n^H\Xi\mathbf{e}_i|^2 - |\mathbf{e}_n^H\Sigma^{-1}\mathbf{e}_i|^2$, $\varphi_{n,m} \triangleq |\mathbf{e}_n^H\Xi\mathbf{f}_m|^2 - |\mathbf{e}_n^H\Sigma^{-1}\mathbf{f}_m|^2$, $\phi_n \triangleq |\mathbf{e}_n^H\Xi\mathbf{h}|^2$, and $\varphi_n \triangleq \mathbf{e}_n^H[\Xi - \Sigma^{-1}]\mathbf{e}_n$.

Similarly, (25) is approximated as

$$t\mathbf{f}_m^H\Xi\Delta\mathbf{D}\Xi\mathbf{f}_m - t\mathbf{f}_m^H\Sigma^{-1}\Delta\mathbf{D}\Sigma^{-1}\mathbf{f}_m + t \sum_{j=1}^N \Delta\lambda_j\vartheta_{m,j} \\ + t \sum_{s=1}^M \Delta\mu_s\vartheta_{m,s} + t\Delta w\alpha_m - \Delta\mu_m\mu_m^{-2} \\ - t\Delta\eta_2I_m = t\vartheta_m - \mu_m^{-1} + t\eta_2I_m, \forall m \quad (28)$$

where $\vartheta_{m,j} \triangleq |\mathbf{f}_m^H\Xi\mathbf{e}_j|^2 - |\mathbf{f}_m^H\Sigma^{-1}\mathbf{e}_j|^2$, $\vartheta_{m,s} \triangleq |\mathbf{f}_m^H\Xi\mathbf{f}_s|^2 - |\mathbf{f}_m^H\Sigma^{-1}\mathbf{f}_s|^2$, $\alpha_m \triangleq |\mathbf{f}_m^H\Xi\mathbf{h}|^2$, and $\vartheta_m \triangleq \mathbf{f}_m^H[\Xi - \Sigma^{-1}]\mathbf{f}_m$.

Similarly, we can approximate (26) as

$$t\mathbf{D}\Xi\Delta\mathbf{D}\Xi\mathbf{D} - t\mathbf{D}\Sigma^{-1}\Delta\mathbf{D}\Sigma^{-1}\mathbf{D} + t \sum_{n=1}^N \Delta\lambda_n\mathbf{D}\Xi\mathbf{e}_n\mathbf{e}_n^H\mathbf{D} \\ + t \sum_{m=1}^M \Delta\mu_m\mathbf{D}\Xi\mathbf{f}_m\mathbf{f}_m^H\mathbf{D} + t\Delta w\mathbf{D}\Xi\mathbf{D} - \Delta\mathbf{D} - t\Delta\eta_2\xi\mathbf{D}^2 \\ = t\mathbf{D}[\Xi - \Sigma^{-1}]\mathbf{D} - \mathbf{D} + t\eta_2\xi\mathbf{D}^2 \quad (29)$$

where $\dot{\Xi}_n \triangleq \Xi\mathbf{e}_n\mathbf{e}_n^H\Xi - \Sigma^{-1}\mathbf{e}_n\mathbf{e}_n^H\Sigma^{-1}$, $\dot{\Xi}_m \triangleq \Xi\mathbf{f}_m\mathbf{f}_m^H\Xi - \Sigma^{-1}\mathbf{f}_m\mathbf{f}_m^H\Sigma^{-1}$, and $\dot{\Xi} \triangleq \Xi\mathbf{h}\mathbf{h}^H\Xi$. Finally, from (19b), (19c), and (19g), we have

$$\Delta\dot{\mathbf{U}} + t\Delta\eta_1\dot{\mathbf{U}}^2 = \dot{\mathbf{U}} - t\eta_1\dot{\mathbf{U}}^2 \quad (30)$$

$$\Delta w + \text{tr}(\Delta\dot{\mathbf{U}}) = P - w - \text{tr}(\dot{\mathbf{U}}) \quad (31)$$

$$\xi\text{tr}(\Delta\mathbf{D}) + \mathbf{p}^T\Delta\psi = P - \xi\text{tr}(\mathbf{D}) - \mathbf{p}^T\psi. \quad (32)$$

At each Newton step, we can transform (23), (27), (28), (29), (30), (31), and (32) into a system of linear equations (i.e., vectorize \mathbf{D} and $\dot{\mathbf{U}}$ as a vector of length $N(N+1)/2$ and $N(N-1)/2$, respectively) to find the optimization variables. However, such a generic method using elimination requires complexity of $\mathcal{O}(N^6)$. In what follows, we apply a block elimination method to find the optimal primal and dual variables which results a low-complexity [33] as

$$\Delta\dot{\mathbf{U}} = \Psi^{(0)} + \Delta\eta_1\Psi^{(1)}. \quad (33)$$

$$\Delta\mathbf{D} = \Theta^{(0)} + \sum_{n=1}^N \Delta\lambda_n\Theta^{(n)} + \sum_{m=1}^M \Delta\mu_m\Theta^{(N+m)} \\ + \Delta w\Theta^{(N+M+1)} + \Delta\eta_2\Theta^{(N+M+2)}. \quad (34)$$

Substituting (33) into (30), we have

$$\Psi^{(0)} = \dot{\mathbf{U}} - t\eta_1\dot{\mathbf{U}}^2 \\ \Psi^{(1)} = -t\dot{\mathbf{U}}^2. \quad (35)$$

The complexity of computing the multiplication of two matrices in (35) is of the order $\mathcal{O}((N-1)^3)$. By doing so, we substitute (34) into (29) yielding a system of $(N+M+3)$

generalized Sylvester equations

$$\begin{aligned}
& t\mathbf{D}\Sigma^{-1}\Theta^{(0)}\Sigma^{-1}\mathbf{D} - t\mathbf{D}\Xi\Theta^{(0)}\Xi\mathbf{D} + \Theta^{(0)} \\
& \quad = -t\mathbf{D}[\Xi - \Sigma^{-1} + \eta_2\xi\mathbf{I}]\mathbf{D} + \mathbf{D} \\
& t\mathbf{D}\Sigma^{-1}\Theta^{(n)}\Sigma^{-1}\mathbf{D} - t\mathbf{D}\Xi\Theta^{(n)}\Xi\mathbf{D} + \Theta^{(n)} \\
& \quad = t\mathbf{D}\dot{\Xi}_n\mathbf{D}, \forall n \\
& t\mathbf{D}\Sigma^{-1}\Theta^{(N+m)}\Sigma^{-1}\mathbf{D} - t\mathbf{D}\Xi\Theta^{(N+m)}\Xi\mathbf{D} \\
& \quad + \Theta^{(N+m)} = t\mathbf{D}\hat{\Xi}_m\mathbf{D}, \forall m \\
& t\mathbf{D}\Sigma^{-1}\Theta^{(N+M+1)}\Sigma^{-1}\mathbf{D} - t\mathbf{D}\Xi\Theta^{(N+M+1)}\Xi\mathbf{D} \\
& \quad + \Theta^{(N+M+1)} = t\mathbf{D}\bar{\Xi}\mathbf{D} \\
& t\mathbf{D}\Sigma^{-1}\Theta^{(N+M+2)}\Sigma^{-1}\mathbf{D} - t\mathbf{D}\Xi\Theta^{(N+M+2)}\Xi\mathbf{D} \\
& \quad + \Theta^{(N+M+2)} = -t\xi\mathbf{D}^2.
\end{aligned} \tag{36}$$

The numerical methods to solve the generalized Sylvester equations (36) with complexity $\mathcal{O}(N^3)$ can be found in [35]. Thus, it can be said that the complexity of solving (35) and (36) is much lower than that of solving a system of linear equations by a generic method.

To calculate Δw , $\{\Delta\lambda_n\}$, $\{\Delta\mu_m\}$, and $\Delta\eta_i, i \in \{1, 2\}$, at each Newton step, we aim to stack (23), (27), (28), (31), and (32) into a system of linear equations. To do this, we substitute (34) into (23), which leads to

$$\begin{aligned}
\gamma_w\Delta w + \sum_{n=1}^N \gamma_n\Delta\lambda_n + \sum_{m=1}^M \tilde{\gamma}_m\Delta\mu_m \\
+ tw^2\Delta\eta_1 + \gamma_0\Delta\eta_2 = \gamma_0 \tag{37}
\end{aligned}$$

where $\gamma_w, \gamma_n, \gamma_m, \gamma_{\eta_2}$, and γ_0 are, respectively, given as follows

$$\begin{aligned}
\gamma_w &= t\mathbf{h}^H\Xi\Theta^{(N+M+1)}\Xi\mathbf{h}w^2 + t\mathbf{h}^H\Xi\mathbf{h}\mathbf{h}^H\Xi\mathbf{h}w^2 + 1 \\
\gamma_n &= t\mathbf{h}^H\Xi\Theta^{(n)}\Xi\mathbf{h}w^2 + t\mathbf{h}^H\Xi\mathbf{e}_n\mathbf{e}_n^H\Xi\mathbf{h}w^2 \\
\tilde{\gamma}_m &= t\mathbf{h}^H\Xi\Theta^{(N+m)}\Xi\mathbf{h}w^2 + t\mathbf{h}^H\Xi\mathbf{F}_m\Xi\mathbf{h}w^2 \\
\gamma_{\eta_2} &= t\mathbf{h}^H\Xi\Theta^{(N+M+2)}\Xi\mathbf{h}w^2 \\
\gamma_0 &= t\mathbf{h}^H\Xi\mathbf{h}w^2 + w - tw^2\eta_1 - t\mathbf{h}^H\Xi\Theta^{(0)}\Xi\mathbf{h}w^2. \tag{38}
\end{aligned}$$

Similarly to (37), (27) can then be rewritten as

$$\begin{aligned}
\tilde{\phi}_n\Delta w + \sum_{i=1}^N \tilde{\varphi}_{n,i}\Delta\lambda_i - \lambda_n^{-2}\Delta\lambda_n \\
+ \sum_{m=1}^M \tilde{\varphi}_{n,m}\Delta\mu_m + \theta_n\Delta\eta_2 = \tilde{\varphi}_n, \forall n \tag{39}
\end{aligned}$$

where $\tilde{\phi}_n, \tilde{\varphi}_{n,i}, \tilde{\varphi}_{n,m}, \theta_n$, and $\tilde{\varphi}_n$ are, respectively, defined by

$$\begin{aligned}
\tilde{\phi}_n &= t\mathbf{e}_n^H\Xi\Theta^{(N+M+1)}\Xi\mathbf{e}_n \\
& \quad - t\mathbf{e}_n^H\Sigma^{-1}\Theta^{(N+M+1)}\Sigma^{-1}\mathbf{e}_n + t\phi_n \\
\tilde{\varphi}_{n,i} &= t\mathbf{e}_n^H\Xi\Theta^{(i)}\Xi\mathbf{e}_n - t\mathbf{e}_n^H\Sigma^{-1}\Theta^{(i)}\Sigma^{-1}\mathbf{e}_n \\
& \quad + t\varphi_{n,i} \\
\tilde{\varphi}_{n,m} &= t\mathbf{e}_n^H\Xi\Theta^{(N+m)}\Xi\mathbf{e}_n \\
& \quad - t\mathbf{e}_n^H\Sigma^{-1}\Theta^{(N+m)}\Sigma^{-1}\mathbf{e}_n + t\varphi_{n,m}
\end{aligned}$$

$$\begin{aligned}
\theta_n &= t\mathbf{e}_n^H\Xi\Theta^{(N+M+2)}\Xi\mathbf{e}_n \\
& \quad - t\mathbf{e}_n^H\Sigma^{-1}\Theta^{(N+M+2)}\Sigma^{-1}\mathbf{e}_n - tP_n \\
\tilde{\varphi}_n &= t\varphi_n - \lambda_n^{-1} + t\eta_2P_n \\
& \quad - t\mathbf{e}_n^H\Xi\Theta^{(0)}\Xi\mathbf{e}_n + t\mathbf{e}_n^H\Sigma^{-1}\Theta^{(0)}\Sigma^{-1}\mathbf{e}_n.
\end{aligned} \tag{40}$$

Next, (28) is reformulated as

$$\begin{aligned}
\tilde{\alpha}_m\Delta w + \sum_{j=1}^N \tilde{\vartheta}_{m,j}\Delta\lambda_j + \sum_{s=1}^M \tilde{\vartheta}_{m,s}\Delta\mu_s \\
- \mu_m^{-2}\Delta\mu_m + \tilde{\theta}_m\Delta\eta_2 = \tilde{\vartheta}_m, \forall m \tag{41}
\end{aligned}$$

where $\tilde{\alpha}_m, \tilde{\vartheta}_{m,j}, \tilde{\vartheta}_{m,s}, \tilde{\theta}_m$, and $\tilde{\vartheta}_m$ are, respectively, given as

$$\begin{aligned}
\tilde{\alpha}_m &= t\mathbf{f}_m^H\Xi\Theta^{(N+M+1)}\Xi\mathbf{f}_m \\
& \quad - t\mathbf{f}_m^H\Sigma^{-1}\Theta^{(N+M+1)}\Sigma^{-1}\mathbf{f}_m + t\alpha_m \\
\tilde{\vartheta}_{m,j} &= t\mathbf{f}_m^H\Xi\Theta^{(j)}\Xi\mathbf{f}_m - t\mathbf{f}_m^H\Sigma^{-1}\Theta^{(j)}\Sigma^{-1}\mathbf{f}_m \\
& \quad + t\vartheta_{m,j} \\
\tilde{\vartheta}_{m,s} &= t\mathbf{f}_m^H\Xi\Theta^{(N+s)}\Xi\mathbf{f}_m \\
& \quad - t\mathbf{f}_m^H\Sigma^{-1}\Theta^{(N+s)}\Sigma^{-1}\mathbf{f}_m + t\vartheta_{m,s} \\
\tilde{\theta}_m &= t\mathbf{f}_m^H\Xi\Theta^{(N+M+2)}\Xi\mathbf{f}_m \\
& \quad - t\mathbf{f}_m^H\Sigma^{-1}\Theta^{(N+M+2)}\Sigma^{-1}\mathbf{f}_m - tI_m \\
\tilde{\vartheta}_m &= t\vartheta_m - \mu_m^{-1} + t\eta_2PI_m - t\mathbf{f}_m^H\Xi\Theta^{(0)}\Xi\mathbf{f}_m \\
& \quad + t\mathbf{f}_m^H\Sigma^{-1}\Theta^{(0)}\Sigma^{-1}\mathbf{f}_m.
\end{aligned} \tag{42}$$

Finally, substituting (33) and (34) into (31) and (32), respectively, yields

$$\Delta w + \text{tr}(\Psi^{(1)})\Delta\eta_1 = P - w - \text{tr}(\dot{\mathbf{U}} + \Psi^{(0)}) \tag{43}$$

$$\chi_w\Delta w + \sum_{n=1}^N \chi_n\Delta\lambda_n + \sum_{m=1}^M \tilde{\chi}_m\Delta\mu_m + \chi_{\eta_2}\Delta\eta_2 = \chi_0 \tag{44}$$

where $\chi_w = \xi\text{tr}(\Theta^{(N+M+1)})$, $\chi_n = \xi\text{tr}(\Theta^{(n)}) + P_n$, $\tilde{\chi}_m = \xi\text{tr}(\Theta^{(N+m)}) + I_m$, $\chi_{\eta_2} = \xi\text{tr}(\Theta^{(N+M+2)})$, and $\chi_0 = P - \mathbf{p}^T\psi - \xi\text{tr}(\mathbf{D} + \Theta^{(0)})$. Let us define $\Delta\mathbf{x} \triangleq [\Delta w \Delta\lambda^T \Delta\mu^T \Delta\eta_1 \Delta\eta_2]^T$. A system of linear equations can be derived by stacking (37), (39), (41), (43), and (44) as

$$\mathbf{A}\Delta\mathbf{x} = \mathbf{b} \tag{45}$$

where $b_1 = \gamma_0, b_{n+1} = \tilde{\varphi}_n$ for $n = 1, \dots, N$, $b_{N+1+m} = \tilde{\vartheta}_m$ for $m = 1, \dots, M$, $b_{N+M+2} = P - w - \text{tr}(\dot{\mathbf{U}} + \Psi^{(0)})$, and $b_{N+M+3} = \chi_0$. Next, the matrix $\mathbf{A} \in \mathbb{C}^{(N+M+3) \times (N+M+3)}$ is given at the top of the next page, where $\delta_{i,j}$ denotes the Kronecker's function, i.e., $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. We point out that the complexity of solving (45) (i.e., the inverse of matrix \mathbf{A}) is of the order $\mathcal{O}((N+M+1)^3)$ [33].

In order to apply the barrier method, we need to compute the residual norm of $w, \dot{\mathbf{U}}, \mathbf{D}, \lambda, \mu$, and $\{\eta_i\}$ used in the

$$\mathbf{A} = \begin{bmatrix} \gamma_w & \cdots & \gamma_n & \cdots & \tilde{\gamma}_m & \cdots & tw^2 & \gamma_{\eta_2} \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ \tilde{\phi}_n & \cdots & \tilde{\varphi}_{n,i} - \lambda_n^{-2} \delta_{n,i} & \cdots & \tilde{\varphi}_{n,m} & \cdots & 0 & \theta_n \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ \tilde{\alpha}_m & \cdots & \tilde{\vartheta}_{m,j} & \cdots & \tilde{\vartheta}_{m,s} - \mu_m^{-2} \delta_{m,s} & \cdots & 0 & \tilde{\theta}_m \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 & \cdots & \text{tr}(\Psi^{(1)}) & 0 \\ \chi_w & \cdots & \chi_n & \cdots & \tilde{\chi}_m & \cdots & 0 & \chi_{\eta_2} \end{bmatrix}.$$

backtracking line search procedure which is defined as [29]

$$\begin{aligned} r(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \{\eta_i\}) = & \\ & |\mathbf{h}^H(\boldsymbol{\Sigma} + \mathbf{h}\mathbf{w}\mathbf{h}^H)^{-1}\mathbf{h} + \frac{1}{t}w^{-1} - \eta_1| \\ & + \|\frac{1}{t}\dot{\mathbf{U}}^{-1} - \eta_1\mathbf{I}\|_F + \|(\boldsymbol{\Sigma} + \mathbf{h}\mathbf{w}\mathbf{h}^H)^{-1} \\ & - \boldsymbol{\Sigma}^{-1} - \frac{1}{t}\mathbf{D}^{-1} + \eta_2\xi\mathbf{I}\|_F + \|\boldsymbol{\pi}\|_2 + \|\boldsymbol{\varpi}\|_2 \\ & + |P - w - \text{tr}(\dot{\mathbf{U}})| + |P - \xi\text{tr}(\mathbf{D}) - \mathbf{p}^T\boldsymbol{\psi}| \end{aligned} \quad (46)$$

where $\boldsymbol{\pi} \in \mathbb{C}^{N \times 1}$ and $\boldsymbol{\varpi} \in \mathbb{C}^{M \times 1}$ are, respectively, defined by

$$\begin{aligned} \pi_n &\triangleq \mathbf{e}_n^H [(\boldsymbol{\Sigma} + \mathbf{h}\mathbf{w}\mathbf{h}^H)^{-1} - \boldsymbol{\Sigma}^{-1}] \mathbf{e}_n - \frac{1}{t}\lambda_n^{-1} + \eta_2 P_n, \forall n \\ \varpi_m &\triangleq \mathbf{f}_m^H [(\boldsymbol{\Sigma} + \mathbf{h}\mathbf{w}\mathbf{h}^H)^{-1} - \boldsymbol{\Sigma}^{-1}] \mathbf{f}_m - \frac{1}{t}\mu_m^{-1} + \eta_2 I_m, \forall m. \end{aligned}$$

The proposed numerical algorithm based on the barrier method to solve (13) is summarized in Algorithm 1. The backtracking line search procedure in line 12 stops when the residual norms in (46) is less than a given accuracy, i.e., the tolerance ϵ .

B. Optimization over Γ_{tol}

P7 can be rewritten by considering the optimization over Γ_{tol} as

$$\max_{\Gamma_{\text{tol}} > 0} f(\Gamma_{\text{tol}}) \quad (47)$$

where $f(\Gamma_{\text{tol}})$ is defined as

$$\begin{aligned} f(\Gamma_{\text{tol}}) = & \\ \min_{\boldsymbol{\psi} \geq 0, \mathbf{D} \succeq \mathbf{0}} \max_{w \geq 0, \dot{\mathbf{U}} \succeq \mathbf{0}} & \log \frac{|\boldsymbol{\Sigma} + \mathbf{h}\mathbf{w}\mathbf{h}^H|}{|\boldsymbol{\Sigma}|} - \log(1 + \Gamma_{\text{tol}}) \\ \text{s. t. } & w + \text{tr}(\dot{\mathbf{U}}) = P \\ & \xi \text{tr}(\mathbf{D}) + \mathbf{p}^T\boldsymbol{\psi} = P. \end{aligned} \quad (48)$$

Since the objective function is concave with respect to Γ_{tol} , a conventional method to find the optimal solution of Γ_{tol} is based on one dimensional search [14]. However, the major computational complexity comes from solving (48). Therefore, one dimensional search method to seeking a saddle point of Γ_{tol} may not be efficient since one dimensional search often shows slow convergence. In this section, we propose an efficient method to find the global optimal Γ_{tol}^* which greatly reduces the complexity. To do this, we consider the following

Algorithm 1 The proposed numerical algorithm to solve (13)

Initialization: $w := 1, \dot{\mathbf{U}} := \mathbf{I}_{N-1}, \mathbf{D} := \mathbf{I}_N, \boldsymbol{\lambda} := \mathbf{1}, \boldsymbol{\mu} := \mathbf{1}, \{\eta_i\} := 0, t := t_0$, and ℓ and tolerance $\epsilon > 0$

- 1: **repeat** {Outer iteration}
- 2: **repeat** {Inner iteration (centering step)}
- 3: Solve (35) to find $\Psi^{(i)}$ for $0 \leq i \leq 1$.
- 4: Solve (36) to find $\Theta^{(i)}$ for $0 \leq i \leq N + M + 2$.
- 5: Solve (45) to find $\Delta w, \Delta \boldsymbol{\lambda}, \Delta \boldsymbol{\mu}$, and $\{\Delta \eta_i\}$.
- 6: Backtracking line search on $r(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \{\eta_i\})$:
- 7: $s = 1$
- 8: **while** $r(w + s\Delta w, \dot{\mathbf{U}} + s\Delta \dot{\mathbf{U}}, \mathbf{D} + s\Delta \mathbf{D}, \boldsymbol{\lambda} + s\Delta \boldsymbol{\lambda}, \boldsymbol{\mu} + s\Delta \boldsymbol{\mu}, \{\eta_i\} + s\{\Delta \eta_i\}) > (1 - \alpha s)r(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \{\eta_i\})$ or $\dot{\mathbf{U}} + s\Delta \dot{\mathbf{U}} \preceq \mathbf{0}$ and $\mathbf{D} + s\Delta \mathbf{D} \preceq \mathbf{0}$ **do**
- 9: $s = \beta s$
- 10: **end while**
- 11: Update primal and dual variables: $w := w + s\Delta w, \dot{\mathbf{U}} := \dot{\mathbf{U}} + s\Delta \dot{\mathbf{U}}, \mathbf{D} := \mathbf{D} + s\Delta \mathbf{D}, \boldsymbol{\lambda} := \boldsymbol{\lambda} + s\Delta \boldsymbol{\lambda}, \boldsymbol{\mu} := \boldsymbol{\mu} + s\Delta \boldsymbol{\mu}$, and $\{\eta_i\} := \{\eta_i\} + s\{\Delta \eta_i\}$.
- 12: **until** $r(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \{\eta_i\}) < \epsilon$
- 13: Increase t : $t = \ell t$.
- 14: **until** t is sufficiently large to tolerate the duality gap.

equivalent problem for a given set of $(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu})$

$$\begin{aligned} \min_{\Gamma_{\text{tol}} > 0} \quad & h(\Gamma_{\text{tol}}) \triangleq \log(1 + \Gamma_{\text{tol}}) \\ \text{s. t. } \quad & -\text{tr}(\tilde{\mathbf{V}}^H \mathbf{D} \tilde{\mathbf{V}} \bar{\mathbf{U}}) \Gamma_{\text{tol}} = P - w - \text{tr}(\tilde{\boldsymbol{\Omega}} \bar{\mathbf{U}}) \\ & \tilde{\xi} \text{tr}(\mathbf{D}) \Gamma_{\text{tol}} = P - \mathbf{p}^T \boldsymbol{\psi} \end{aligned} \quad (49)$$

where $\tilde{\boldsymbol{\Omega}} \triangleq \tilde{\mathbf{V}}^H (\text{diag}(\boldsymbol{\lambda}) + \sum_{m=1}^M \mu_m \mathbf{F}_m) \tilde{\mathbf{V}}$ and $\tilde{\xi} \triangleq \Phi_N^{-1}(1 - \kappa^{1/K})$. With implicit constraint $\Gamma_{\text{tol}} > 0$, we can derive the simpler form of the objective function in (49) without affecting optimality, as follows:

$$\tilde{h}(\Gamma_{\text{tol}}) \triangleq h(\Gamma_{\text{tol}} - 1) = \log(\Gamma_{\text{tol}}). \quad (50)$$

The Lagrangian function of (49) can be defined as [33]

$$\mathcal{L}(\Gamma_{\text{tol}}, \mathbf{v}) = \tilde{h}(\Gamma_{\text{tol}}) + \mathbf{v}^T \mathbf{a} \Gamma_{\text{tol}} - \boldsymbol{\rho}^T \mathbf{v} \quad (51)$$

where $\boldsymbol{\rho} \triangleq [P - w - \text{tr}(\tilde{\boldsymbol{\Omega}} \bar{\mathbf{U}}) \quad P - \mathbf{p}^T \boldsymbol{\psi}]^T$, $\mathbf{a} \triangleq [-\text{tr}(\tilde{\mathbf{V}}^H \mathbf{D} \tilde{\mathbf{V}} \bar{\mathbf{U}}), \quad \tilde{\xi} \text{tr}(\mathbf{D})]^T$, and $\mathbf{v} \triangleq [v_1 \quad v_2]^T$ with v_1 and v_2 being the dual variables related to the constraints in (49). Then, the solution of (49) can be found by solving the dual problem which is presented in the following theorem

Theorem 2: The dual problem of (49) is given by

$$\max_{\mathbf{v}} g(\mathbf{v}) = -\boldsymbol{\rho}^T \mathbf{v} - \log(-\mathbf{a}^T \mathbf{v}) - 1 \quad (52)$$

with implicit constraint $\mathbf{a}^T \mathbf{v} < 0$.

Proof: See Appendix D. ■

Corollary 1: The optimal solution Γ_{tol} can be obtained from that of (52) as

$$\Gamma_{\text{tol}} = -\frac{1}{\mathbf{v}^T \mathbf{a}}. \quad (53)$$

Proof: Once the optimal dual variable \mathbf{v}^* is found, we can obtain the optimal solution of Γ_{tol} from (49) by solving the KKT condition in (51) as

$$\nabla_{\Gamma_{\text{tol}}} \tilde{h}(\Gamma_{\text{tol}}) + \mathbf{v}^T \mathbf{a} = \Gamma_{\text{tol}}^{-1} + \mathbf{v}^T \mathbf{a} = 0. \quad (54)$$

This completes the proof. ■

We now apply Broyden-Fletcher-Goldfarb-Shanno algorithm [36] to solve the unconstrained optimization problem (52). The BFGS method is a member of the quasi-Newton methods which often shows superior convergence rate. A standard method for solving the log concave function in (52) is to consider the following equivalent optimization problem

$$\min_{\mathbf{v}} \tilde{g}(\mathbf{v}) = -g(\mathbf{v}). \quad (55)$$

The BFGS algorithm can be described as

1. Compute $\Delta \mathbf{v}$ by $\Delta \mathbf{v} = -\mathbf{P} \nabla_{\mathbf{v}} \tilde{g}(\mathbf{v})$, where $\nabla_{\mathbf{v}} \tilde{g}(\mathbf{v}) = \boldsymbol{\rho} - \mathbf{a}/(\mathbf{a}^T \mathbf{v})$.
2. If $\|\nabla_{\mathbf{v}} \tilde{g}(\mathbf{v} + \Delta \mathbf{v})\| < \epsilon$ (tolerance), then stop.
3. Update the dual variables: $\mathbf{v} := \mathbf{v} + \Delta \mathbf{v}$.
4. Compute $\phi' = \nabla_{\mathbf{v}} \tilde{g}(\mathbf{v} + \Delta \mathbf{v}) - \nabla_{\mathbf{v}} \tilde{g}(\mathbf{v})$.
5. Update \mathbf{P} as [36]

$$\mathbf{P} := \mathbf{P} + \frac{(\Delta \mathbf{v}^T \phi' + (\phi')^T \mathbf{P} \phi') (\Delta \mathbf{v} \Delta \mathbf{v}^T)}{(\Delta \mathbf{v}^T \phi')^2} - \frac{\mathbf{P} \phi' \Delta \mathbf{v}^T + \Delta \mathbf{v} (\phi')^T \mathbf{P}}{\Delta \mathbf{v}^T \phi'}. \quad (56)$$

6. Go back to Step 1.

However, the cost of computation of these steps is relatively high when the first gradient on \mathbf{v} is computed many times. To reduce the computational complexity, we choose $\phi = -\nabla \tilde{g}(\mathbf{v})$ and store it for the next iteration. The overall iterative algorithm based on the BFGS method, but customized to our problem to solve (47) is given in Algorithm 2 and it can be summarized as follows. For a given \mathbf{v} , Γ_{tol} in (53) is computed, (48) (or equivalently (13)) is solved based on Algorithm 1 to obtain $(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu})$; then update \mathbf{v} from line 5 until $\|\phi\|$ is below a specified accuracy level (i.e., the tolerance ϵ).

C. Complexity Comparison

The complexity of the proposed method mainly comes from Algorithm 1. In particular, solving (35), (36), and (45) requires a complexity of $\mathcal{O}((N-1)^3)$, $\mathcal{O}(N^3)$, and $\mathcal{O}((N+M+1)^3)$, respectively, while a generic method has complexity of $\mathcal{O}(N^6)$. That is to say, the proposed algorithm requires much less complexity, compared to the generic method. As

Algorithm 2 Main algorithm to solve (47)

Initialization: $\mathbf{v} := \mathbf{1}$, $\phi := -\nabla_{\mathbf{v}} \tilde{g}(\mathbf{v})$, $\phi' := \mathbf{0}$, $\mathbf{P} := \mathbf{I}$, $s' := 1$, and tolerance $\epsilon > 0$

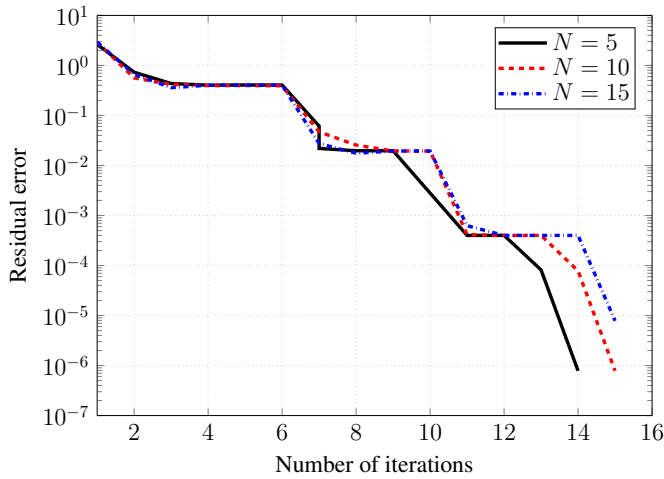
- 1: **loop**
- 2: Solve (48) based on Algorithm 1 to obtain $(w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.
- 3: Compute $\phi := \phi - \phi'$ and $\Delta \mathbf{v} = \mathbf{P} \phi$.
- 4: Stop, if $\|\phi\| < \epsilon$.
- 5: Update \mathbf{v} , ϕ' , and \mathbf{P} in the strict order:
 $\mathbf{v} := \mathbf{v} + s' \Delta \mathbf{v}$
 $\phi' := \phi + \nabla_{\mathbf{v}} \tilde{g}(\mathbf{v})$
 \mathbf{P} as the result in (56).
- 6: **end loop**
- 7: **Output:** $(\mathbf{v}, w, \dot{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu})$

mentioned in Section III, a semidefinite program (SDP) can be applied to solve the relaxed problem of **P6**, i.e., using a specific interior point method which is called the primal-dual path following method [37]. However, such a method has complexity of $\mathcal{O}((2N-1)^4)$ per iteration, which is higher than that of our proposed method, especially when N becomes large.

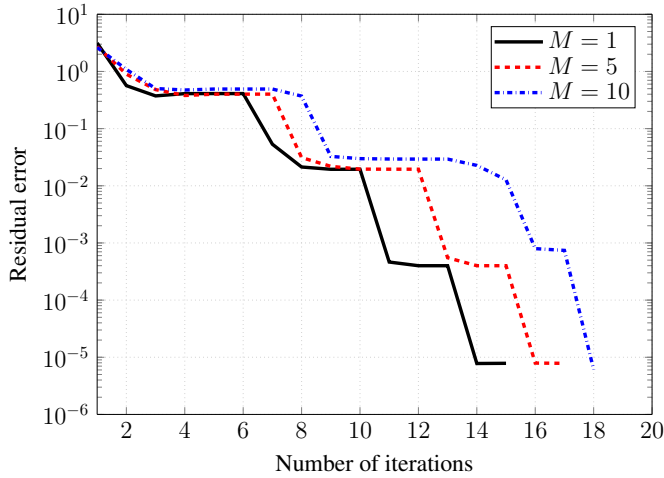
V. NUMERICAL RESULTS

In this section, we provide numerical results to validate the performance of the proposed optimal approach. The entries of the channel vectors are all generated as independent circularly symmetric complex Gaussian (CSCG) random variables with zero-mean and unit variance. To guarantee secure communication, we set the probability $\kappa = 0.99$. For simplicity, we assume that the interference thresholds at the PUs are equal, i.e., $I_m = I$ for all m , and the number of Eves is fixed to $K = 3$. The resulting power constraint for each antenna is $P_n = P/N$ for all n , where P is the total transmit power at the ST. We also compare the performance of the proposed scheme with existing schemes, namely, the “Isotropic JN scheme” [9], [22] and “No JN scheme” [5]. In the “Isotropic JN scheme,” the covariance matrix of the jamming beamforming is chosen as $\mathbf{U} = \frac{p_u \mathbf{V} \mathbf{V}^H}{N-1}$ where the variable p_u is used to control interference to the PUs and Eves. In the “No JN scheme,” the optimization problem with no JN is considered as a benchmark where the optimal solution can be obtained from (47) by setting \mathbf{U} to $\mathbf{0}$. The results obtained from (47) are referred to as the “Optimal JN scheme.” The results of the average secrecy rate are shown by averaging over 1,000 simulation trials.

Fig. 2 depicts the typical convergence behavior of the proposed Algorithm 1 where the network parameters are given in the caption. Particularly, we plot the typical convergence behavior of the proposed barrier method as a function of the number of transmit antennas at the ST in Fig. 2(a), and a function of the number of PUs in Fig. 2(b). The initial values for the primal and dual variables in Algorithm 1 are randomly generated. It is observed that Algorithm 1 exhibits a fast convergence rate, which is slightly sensitive to the network configurations.



(a) Convergence behavior of the proposed Algorithm 1 for different number of transmit antennas at the ST. The number of PUs is $M = 1$.



(b) Convergence behavior of the proposed Algorithm 1 for different number of PUs. The number of transmit antennas at the ST is $N = 5$.

Fig. 2. Convergence behavior of the proposed Algorithm 1, (a) for different number of transmit antennas at the ST, and (b) for different number of PUs. Each curve is obtained for one channel realization. The parameters of Algorithm 1 are as follows. The tolerance is set to $\epsilon = 10^{-5}$. The barrier parameters ℓ and t_0 are set to 1 and 50, respectively. The backtracking line search parameters in Algorithm 1 are set as $\alpha = 0.01$ and $\beta = 0.5$. In this example, we set the network parameters as $P = 10$ dB and $I = 1$ dB.

Fig. 3 shows the convergence rate of the proposed Algorithm 2. We also compare the convergence rate of the proposed algorithm with the *damped* Newton method. The *damped* Newton method can easily be applied to solve the considered problem following similar steps as in [33, Section 9.5.2] which also shows a fast convergence. The initial value of Γ_{tol} is set to 0 dB. As seen, the duality gap of the Algorithm 2 drops fast to minimum when the number of iterations increases, and shows a faster convergence than that of the *damped* Newton method algorithm. This is mainly due to the fact that the BFGS algorithm approximates the inverse of the Hessian matrix by using rank-one updates [36], which means that the Hessian matrix is not computed directly. We recall that this result significantly reduces the complexity of the considered problem since the major computational complexity comes

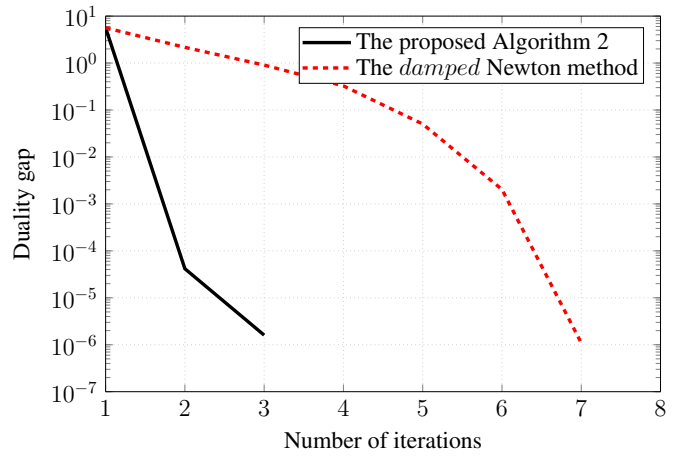


Fig. 3. Convergence behavior of the proposed Algorithm 2. Each curve is obtained for one channel realization. The tolerance ϵ is set to 10^{-5} . The network configuration is $N = 5$, $M = 3$, $P = 10$ dB, and $I = 1$ dB.

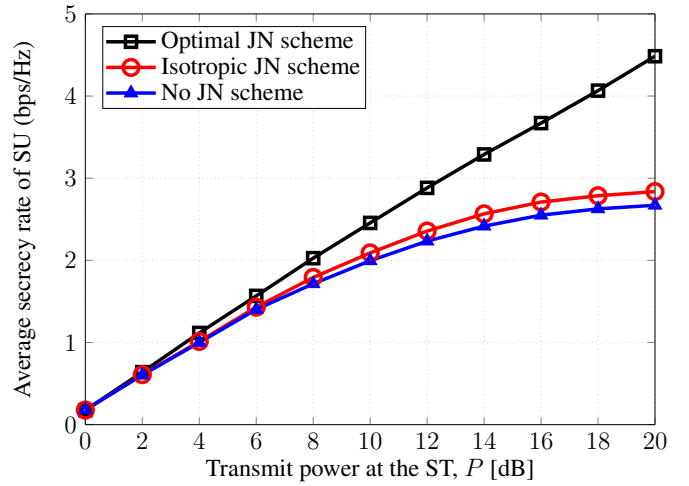
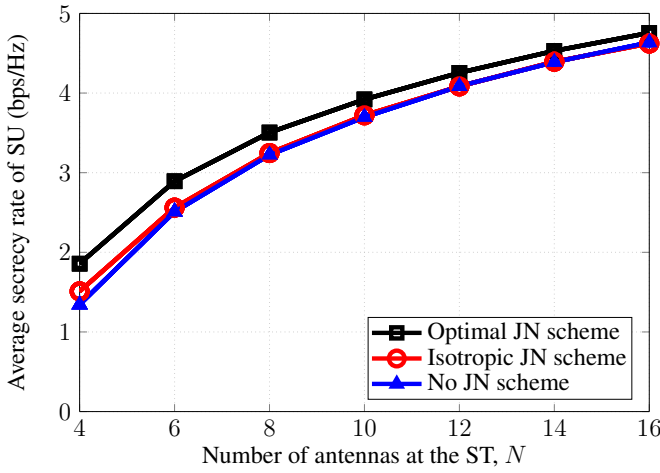


Fig. 4. Average secrecy rate of the secondary system versus the total transmit power at the ST. In this example, we set the network parameters as $I = 5$ dB, $N = 5$, and $M = 3$.

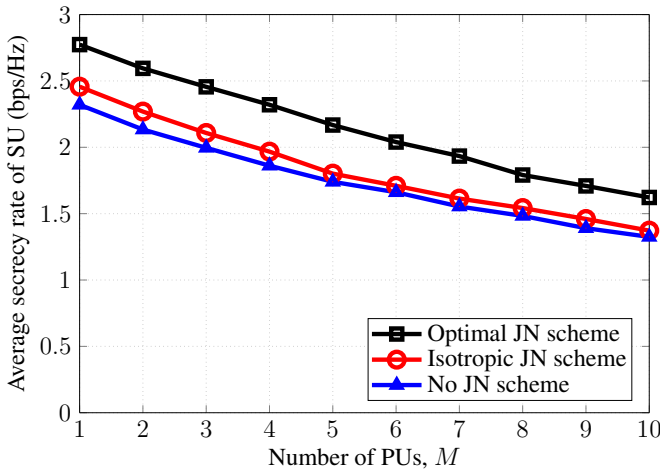
from Algorithm 1.

Fig. 4 illustrates the average secrecy rate of the secondary system versus the transmit power at the ST. As seen, the curves coincide in low power regime. The reason is that in such a case, the ST mainly focuses on maximizing the secrecy rate, as the interference constraints are likely satisfied for all PUs. This indicates that JN may not be necessary in this regime. However, in high power regime, the schemes using JN outperform the scheme with no JN in terms of the secrecy rate. In addition, for the “Optimal JN scheme,” the ST is allowed to transmit with nearly full power, whereas the performance of the “Isotropic JN scheme” and “No JN scheme” tends to saturate. This is because the “Optimal JN scheme” controls the interference to the PUs more efficiently than the other schemes thanks to the optimized transmission.

The average secrecy rate of the secondary system is plotted versus the number of antennas at the ST and the number of PUs in Fig. 5(a) and Fig. 5(b), respectively. From Fig. 5(a), it can be observed that the average secrecy rate of the “Isotropic JN scheme” increases with N and approaches that



(a) Average secrecy rate of the secondary system for different number of transmit antennas at the ST. The number of PUs is $M = 3$.



(b) Average secrecy rate of the secondary system for different number of PUs. The number of transmit antennas at the ST is $N = 5$.

Fig. 5. Average secrecy rate of the secondary system, (a) for different number of transmit antennas at the ST, and (b) for different number of PUs. The network parameters are set to $P = 10$ dB and $I = 5$ dB.

of the “No JN scheme” for large N . This means that for the “Isotropic JN scheme,” we need little or no JN when N is sufficiently large. The “Optimal scheme” still achieves a better performance than the other schemes in all the range of N . As expected, the average secrecy rate is improved as the number of transmit antennas increases in all schemes, since more degrees of freedom are added to the ST. From Fig. 5(b), the performance of the three schemes degrades significantly when the number of PUs increases. The reason is that when the transmit power becomes sufficiently large, the ST needs to avoid transmitting its signals over the spatial space of PUs. Therefore, the degrees of freedom left for the ST are reduced when the number of PUs increases.

In this numerical example, we plot the average secrecy rate of the secondary system for the “Optimal JN scheme” under several different assumptions of sharing equally the resources, i.e., transmit power at the ST and interference thresholds at the PUs. In particular, for total transmit power at the ST, the information and JN beamforming are assumed

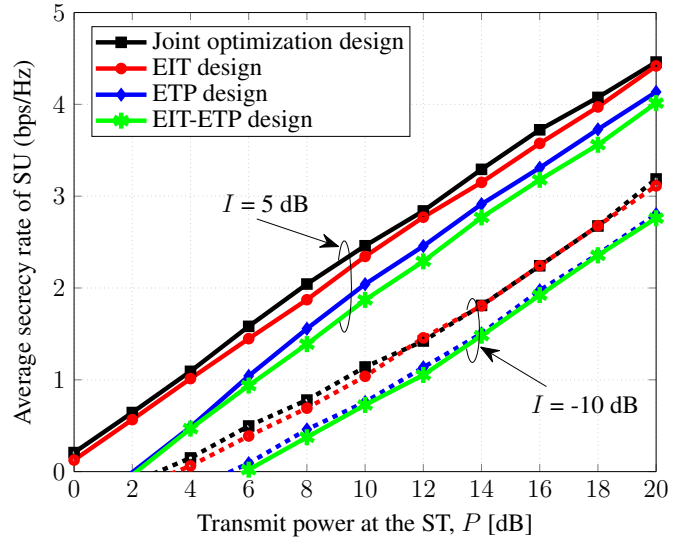


Fig. 6. Average secrecy rate of the secondary system versus the total transmit power at the ST, P , for different interference thresholds at the PUs, I ($I = I_m, m = 1, 2, \dots, M$). The network configuration is $N = 5$ and $M = 3$.

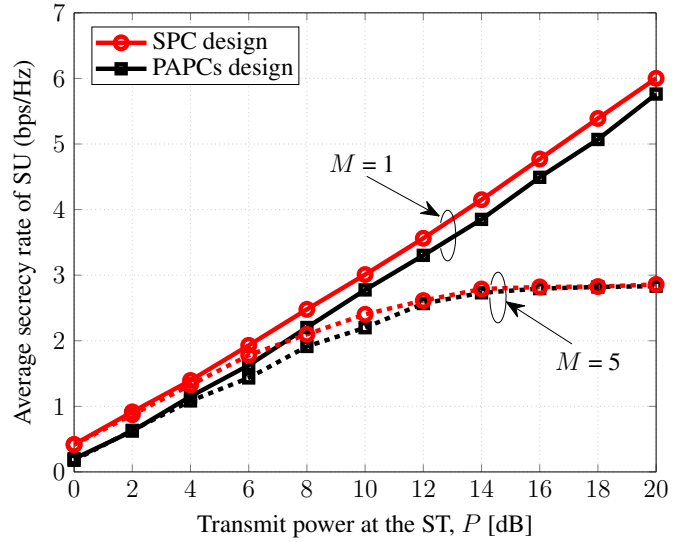


Fig. 7. Average secrecy rate of the secondary system versus the total transmit power at the ST, P , for different number of PUs, M . We set the network parameters as $N = 5$ and $I = 5$ dB.

to share 50% of the power resource, i.e., $[\mathbf{W}]_{n,n} \leq P_n/2$ and $[\mathbf{V}\mathbf{U}\mathbf{V}^H]_{n,n} \leq P_n/2$, which is referred to as equal transmit power (ETP). Likewise, the information and JN beamforming are assumed to share 50% of the interference threshold, i.e., $\text{tr}(\mathbf{F}_m \mathbf{W}) \leq I_m/2$ and $\text{tr}(\mathbf{F}_m \mathbf{U}) \leq I_m/2$, which is referred to as equal interference threshold (EIT). Consequently, we compare the performance of the proposed design with three other suboptimal methods, namely ETP, EIT, and EIT-ETP, and we present the results in Fig. 6. A general observation is that the joint optimization design outperforms the other designs in terms of the secrecy rate of the secondary system, especially when compared to the ETP design. In addition, decreasing the interference threshold I significantly degrades the secrecy rate of the secondary system. The performance gain achieved for higher interference threshold is due to the fact that more

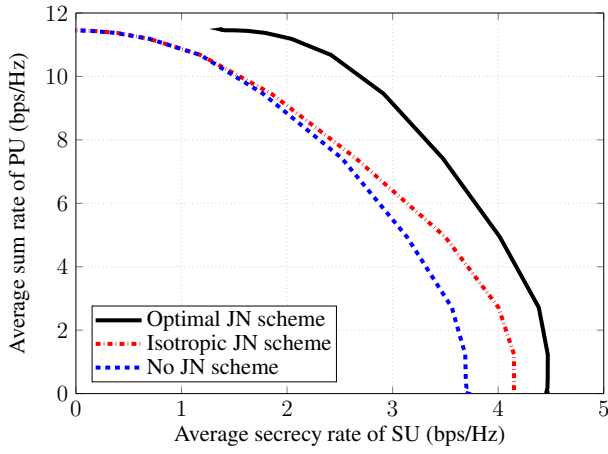


Fig. 8. Average sum rate of the primary system versus average secrecy rate of the secondary system. For both systems, the network parameters are set to $\tilde{N} = N = 5$, $\tilde{P} = P = 15$ dB, and $M = 3$.

transmit power can be used when the interference threshold constraints are set to be high. Interestingly, the secrecy rate of the EIT design approaches the optimal one when the interference threshold is relatively small.

According to Remark 4, we now investigate the effect of two different types of power constraints on the performance of the secondary system, i.e., SPC and PAPCs. As can be seen from Fig. 7, the secondary system achieves a better performance by using the SPC than the PAPCs. The gaps between the two types of power constraints are negligible for a large M and high transmit power at the ST, P . We recall that for large M and high P , the secondary system lacks the degrees of freedom to leverage multiuser diversity.

Finally, the interference from the ST to a PU is considered in several schemes. Thus, it is interesting to investigate how different schemes of interference constraints affect the performance of the primary system. In particular, we consider a primary system where the primary transmitter is equipped with \tilde{N} antennas and adopt zero-forcing beamforming with PAPCs [29]. The interference from the ST is basically treated as background thermal noise at the PUs. The total transmit power at the primary transmitter is limited by \tilde{P} . For this setup, the average achieved sum rate of the primary system is considered as a performance metric to evaluate the effectiveness of each scheme. As can be seen from Fig. 8, the primary system achieves a higher sum rate with the “Optimal scheme” than with the other schemes. The gap between the schemes diminishes for a lower sum rate of the primary system. This is because in such a case, the interference at the PUs is higher, and thus the “Isotropic JN scheme” and “No JN scheme” can manage the interference more effectively. Interestingly, the proposed optimal solution is quite robust even when the sum rate of the primary system is high, whereas the secrecy rate of the “Isotropic JN scheme” and “No JN scheme” drop to zero.

VI. CONCLUSION

We studied the secrecy rate maximization problem of MISO CRN in the presence of multiple passive eavesdroppers and primary users. The problem design is subject to per-antenna

power constraints at the secondary transmitter and interference constraints at the PUs. By assuming the eavesdroppers’ channel i.i.d. Rayleigh fading, we first convert a nonconvex constraint to a linear matrix inequality convex constraint. For an optimal beamforming design, we transformed the problem of secrecy rate maximization into that of finding a saddle point of a minimax program where a rank relaxation method is shown to be tight. We further developed a computationally efficient algorithm based on a barrier method to find the optimal solution of the minimax problem. For the global optimal solution, the BFGS algorithm was employed to solve an unconstrained optimization problem based on a dual problem. Numerical results illustrated the superior convergence behavior of the proposed algorithm, which is robust to the problem size. Through numerical examples, we evaluated the tradeoff between the secrecy rate of the secondary system and the sum rate of the primary system. We concluded that the proposed approach offers a better performance and is quite robust when compared to the existing approaches [5], [9], [22].

APPENDIX A PROOF OF LEMMA 1

The probability in (10b) for the k -th Eve link can be rewritten as

$$\Pr\left(\text{tr}(\mathbf{G}_k(\mathbf{W} - \Gamma_{\text{tol}} \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H)) \leq \Gamma_{\text{tol}}\right). \quad (57)$$

Let $\mathbf{Q} \triangleq \mathbf{W} - \Gamma_{\text{tol}} \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H$. The probability in (57) cannot be computed directly unless specific properties of \mathbf{Q} are satisfied. For $N \times N$ Hermitian matrices \mathbf{G}_k and \mathbf{Q} , the following inequality holds [32]

$$\begin{aligned} \text{tr}(\mathbf{G}_k \mathbf{Q}) &\leq \sum_{i=1}^N \lambda_i(\mathbf{G}_k) \lambda_i(\mathbf{Q}) \\ &\stackrel{(a)}{=} \lambda_{\max}(\mathbf{G}_k) \lambda_{\max}(\mathbf{Q}) \\ &\stackrel{(b)}{=} \text{tr}(\mathbf{G}_k) \lambda_{\max}(\mathbf{Q}) \end{aligned} \quad (58)$$

where $\lambda_i(\mathbf{X})$ denotes the i -th eigenvalue of matrix $\mathbf{X} \in \mathbb{H}^{N \times N}$ and its orders are arranged as $\lambda_{\max}(\mathbf{X}) = \lambda_1(\mathbf{X}) \geq \lambda_2(\mathbf{X}) \geq \dots \geq \lambda_N(\mathbf{X}) = \lambda_{\min}(\mathbf{X})$. In addition, the equalities (a) and (b) in (58) are obtained because \mathbf{G}_k is a rank-one positive semidefinite matrix. Substituting (58) into (57), we have

$$\begin{aligned} &\Pr\left(\text{tr}(\mathbf{G}_k(\mathbf{W} - \Gamma_{\text{tol}} \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H)) \leq \Gamma_{\text{tol}}\right) \\ &\geq \Pr\left(\text{tr}(\mathbf{G}_k) \lambda_{\max}(\mathbf{Q}) \leq \Gamma_{\text{tol}}\right). \end{aligned} \quad (59)$$

Since the channel \mathbf{g}_k , $\forall k$, is modeled as i.i.d. Rayleigh fading, we have

$$\begin{aligned} &\Pr\left(\max_{k \in \mathcal{K}} \frac{\text{tr}(\mathbf{G}_k \mathbf{W})}{\text{tr}(\mathbf{G}_k \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H) + 1} \leq \Gamma_{\text{tol}}\right) \\ &\geq \Pr\left(\text{tr}(\mathbf{G}) \lambda_{\max}(\mathbf{Q}) \leq \Gamma_{\text{tol}}\right) \geq \kappa^{1/K} \\ &\Leftrightarrow \Pr\left(\frac{\lambda_{\max}(\mathbf{Q})}{\Gamma_{\text{tol}}} \geq \frac{1}{\text{tr}(\mathbf{G})}\right) \leq 1 - \kappa^{1/K} \\ &\stackrel{(c)}{\Leftrightarrow} \lambda_{\max}(\mathbf{Q}) \leq \Phi_N^{-1}(1 - \kappa^{1/K}) \Gamma_{\text{tol}} \\ &\Leftrightarrow \mathbf{Q} \preceq \mathbf{I} \left(\Phi_N^{-1}(1 - \kappa^{1/K}) \Gamma_{\text{tol}} \right) \end{aligned} \quad (60)$$

where without loss of generality, we have removed the index of Eves and (c) is obtained similarly to the steps of Lemma 2 in [23]. $\Phi_N^{-1}(\cdot)$ denotes the inverse cumulative distribution function of an inverse central chi-square random variable with $2N$ degrees of freedom, and $\text{tr}(\mathbf{G}) = \text{tr}(|\mathbf{g}|^2)$ is the sum of the squares of N independent Gaussian random variables. This completes the proof.

APPENDIX B PROOF OF LEMMA 2

In this appendix, we prove that the rank of the optimal solution \mathbf{W}^* to (12) is less than or equal to one. The Lagrangian of the relaxed version of (12) can be defined with respect to \mathbf{W} as

$$\begin{aligned} \mathcal{L}(\mathbf{W}, \mathbf{D}, \{\lambda_n\}, \{\mu_m\}, \mathbf{\Lambda}_w) &= \log(1 + \mathbf{h}^H \mathbf{W} \mathbf{h}) - \text{tr}(\mathbf{D} \mathbf{W}) \\ &- \sum_{n=1}^N \lambda_n \text{tr}(\mathbf{W} \mathbf{B}^{(n)}) - \sum_{m=1}^M \mu_m \text{tr}(\mathbf{F}_m \mathbf{W}) + \text{tr}(\mathbf{W} \mathbf{\Lambda}_w) + \Omega \end{aligned} \quad (61)$$

where $\mathbf{B}^{(n)} \triangleq \mathbf{T}^H \mathbf{T}$, $\mathbf{T} = [\mathbf{0}_{n-1}^T \quad 1 \quad \mathbf{0}_{N-n}^T]$, $\mathbf{F}_m \triangleq \mathbf{f}_m \mathbf{f}_m^H$, and Ω denotes the summation of terms that only involve variables not related to the structure of \mathbf{W}^* . $\{\lambda_n \geq 0\}$ and $\{\mu_m \geq 0\}$ are the Lagrange multipliers of P.6 associated with the constraints (12c) and (12d), respectively. The matrices $\mathbf{D} \succeq \mathbf{0}$ and $\mathbf{\Lambda}_w \succeq \mathbf{0}$ are the Lagrange multiplier matrices associated with the constraint (12b) and the positive semidefinite matrix constraint $\mathbf{W} \succeq \mathbf{0}$, respectively.

Let $\mathbf{\Sigma} \triangleq \mathbf{D} + \mathbf{\bar{B}} + \sum_{m=1}^M \mu_m \mathbf{F}_m$, where $\mathbf{\bar{B}} = \text{diag}(\boldsymbol{\lambda})$ and $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_N]^T$. Then (61) can be rewritten as

$$\begin{aligned} \mathcal{L}(\mathbf{W}, \mathbf{\Sigma}, \mathbf{\Lambda}_w) &= \log(1 + \mathbf{h}^H \mathbf{W} \mathbf{h}) - \text{tr}(\mathbf{\Sigma} \mathbf{W}) \\ &+ \text{tr}(\mathbf{W} \mathbf{\Lambda}_w) + \Omega. \end{aligned} \quad (62)$$

The optimality conditions for (62) are directly related to \mathbf{W}^* , which should satisfy the necessary conditions for optimality [33]

$$\begin{aligned} \frac{\mathbf{h} \mathbf{h}^H}{1 + \mathbf{h}^H \mathbf{W}^* \mathbf{h}} - \mathbf{\Sigma}^* + \mathbf{\Lambda}_w^* &= \mathbf{0} \\ \mathbf{\Lambda}_w^* \mathbf{W}^* &= \mathbf{0} \\ \mathbf{W}^* &\succeq \mathbf{0}. \end{aligned} \quad (63)$$

By multiplying the first equation in (63) with \mathbf{W}^* and applying the result to the second equation in (63), we get

$$\frac{\mathbf{h} \mathbf{h}^H}{1 + \mathbf{h}^H \mathbf{W}^* \mathbf{h}} \mathbf{W}^* = \mathbf{\Sigma}^* \mathbf{W}^*. \quad (64)$$

Since $\mathbf{\Sigma}^*$ is invertible,² $\text{rank}(\mathbf{\Sigma}^*) = N$. In addition, $\text{rank}(\mathbf{h} \mathbf{h}^H) \leq 1$, and it follows from (64) that $\text{rank}(\mathbf{W}^*) \leq 1$. Lemma 2 is thus proved.

²We remark that $\{\lambda_n\}$, or \mathbf{D} and $\{\mu_m\}$ must be positive. This is because the total power at the ST should be used up or the interference constraints should be met at optimum, i.e., $\{\lambda_n^*\} > 0$, or $\mathbf{D}^* \succ \mathbf{0}$ and $\{\mu_m^*\} > 0$. In other words, the inequality constraint in (12c), or (12b) and (12d) hold with equality at optimality.

APPENDIX C PROOF OF THEOREM 1

Here, we prove that P7 is the dual problem of the relaxed problem of P6. In particular, we follow the same steps as in [29] while customizing them to our considered problem. The partial Lagrangian function of the relaxed version of P6 can be defined as

$$\begin{aligned} \mathcal{L}(\mathbf{W}, \bar{\mathbf{U}}, \mathbf{D}, \{\lambda_n\}, \{\mu_m\}) &= \log(1 + \mathbf{h}^H \mathbf{W} \mathbf{h}) \\ &- \text{tr}(\mathbf{D}(\mathbf{W} - \Gamma_{\text{tol}} \bar{\mathbf{V}} \bar{\mathbf{U}} \bar{\mathbf{V}}^H - \mathbf{I} \xi)) \\ &- \sum_{n=1}^N \lambda_n \left(\text{tr}(\mathbf{W} \mathbf{B}^{(n)}) + \text{tr}(\bar{\mathbf{U}} \mathbf{E}^{(n)}) - P_n \right) \\ &- \sum_{m=1}^M \mu_m \left(\text{tr}(\mathbf{F}_m \mathbf{W}) + \text{tr}(\bar{\mathbf{F}}_m \bar{\mathbf{U}}) - I_m \right) \end{aligned} \quad (65)$$

where $\mathbf{E}^{(n)} \triangleq \bar{\mathbf{T}}^H \bar{\mathbf{T}}$, $\bar{\mathbf{T}} = [\mathbf{0}_{n-1}^T \quad 1 \quad \mathbf{0}_{N-n}^T]$, and $\bar{\mathbf{F}}_m \triangleq \bar{\mathbf{f}}_m \bar{\mathbf{f}}_m^H$. Next, the dual objective of P6 is given by

$$\mathcal{D}(\mathbf{D}, \{\lambda_n\}, \{\mu_m\}) = \max_{\mathbf{W}, \bar{\mathbf{U}} \succeq \mathbf{0}} \mathcal{L}(\mathbf{W}, \bar{\mathbf{U}}, \mathbf{D}, \{\lambda_n\}, \{\mu_m\}). \quad (66)$$

For a given set $(\mathbf{D}, \{\lambda_n\}, \{\mu_m\})$, we first rewrite the partial Lagrangian function as

$$\begin{aligned} \mathcal{L}(\mathbf{W}, \bar{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= \log(1 + \mathbf{h}^H \mathbf{W} \mathbf{h}) - \text{tr}(\mathbf{\Sigma} \mathbf{W}) \\ &- \text{tr}(\mathbf{\Omega} \bar{\mathbf{U}}) + \xi \text{tr}(\mathbf{D}) + \bar{\mathbf{p}}^T \boldsymbol{\lambda} + \bar{\mathbf{I}}^T \boldsymbol{\mu} \end{aligned} \quad (67)$$

where $\mathbf{\Omega} \triangleq -\Gamma_{\text{tol}} \bar{\mathbf{V}}^H \mathbf{D} \bar{\mathbf{V}} + \sum_{n=1}^N \lambda_n \mathbf{E}^{(n)} + \sum_{m=1}^M \mu_m \bar{\mathbf{F}}_m = \bar{\mathbf{V}}^H \left(-\Gamma_{\text{tol}} \mathbf{D} + \text{diag}(\boldsymbol{\lambda}) + \sum_{m=1}^M \mu_m \mathbf{F}_m \right) \bar{\mathbf{V}}$, $\bar{\mathbf{p}} = [P_1, P_2, \dots, P_N]^T$, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_N]^T$, $\bar{\mathbf{I}} = [I_1, I_2, \dots, I_M]^T$, and $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_M]^T$. Let $\bar{\mathbf{W}} = \mathbf{\Sigma}^{1/2} \mathbf{W} \mathbf{\Sigma}^{1/2}$. Since $\mathbf{\Sigma}$ is invertible, (67) can be rewritten as

$$\begin{aligned} \mathcal{L}(\bar{\mathbf{W}}, \bar{\mathbf{U}}, \mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= \log(1 + \mathbf{h}^H \mathbf{\Sigma}^{-1/2} \bar{\mathbf{W}} \mathbf{\Sigma}^{-1/2} \mathbf{h}) \\ &- \text{tr}(\bar{\mathbf{W}}) - \text{tr}(\mathbf{\Omega} \bar{\mathbf{U}}) + \xi \text{tr}(\mathbf{D}) + \bar{\mathbf{p}}^T \boldsymbol{\lambda} + \bar{\mathbf{I}}^T \boldsymbol{\mu}. \end{aligned} \quad (68)$$

Based on the results in [28, Appendix A], the dual objective in (66) is equivalent to

$$\begin{aligned} \mathcal{D}(\mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= \max_{w \geq 0, \bar{\mathbf{U}} \succeq \mathbf{0}} \log |\mathbf{I} + \mathbf{\Sigma}^{-1/2} \mathbf{h} w \mathbf{h}^H \mathbf{\Sigma}^{-1/2}| \\ &- w - \text{tr}(\mathbf{\Omega} \bar{\mathbf{U}}) + \xi \text{tr}(\mathbf{D}) + \bar{\mathbf{p}}^T \boldsymbol{\lambda} + \bar{\mathbf{I}}^T \boldsymbol{\mu} \end{aligned} \quad (69)$$

where the relationship between $\bar{\mathbf{W}}$ and w is given by

$$\bar{\mathbf{W}} = \frac{\mathbf{\Sigma}^{-1/2} \mathbf{h} w \mathbf{h}^H \mathbf{\Sigma}^{-1/2}}{\mathbf{h}^H \mathbf{\Sigma}^{-1} \mathbf{h}}. \quad (70)$$

Next, we can write $\mathcal{D}(\mathbf{D}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ in a more compact form as

$$\begin{aligned} \mathcal{D}(\mathbf{D}, \boldsymbol{\psi}) &= \max_{w \geq 0, \bar{\mathbf{U}} \succeq \mathbf{0}} \log |\mathbf{I} + \mathbf{\Sigma}^{-1/2} \mathbf{h} w \mathbf{h}^H \mathbf{\Sigma}^{-1/2}| \\ &- w - \text{tr}(\mathbf{\Omega} \bar{\mathbf{U}}) + \xi \text{tr}(\mathbf{D}) + \bar{\mathbf{p}}^T \boldsymbol{\psi} \\ &= \max_{w \geq 0, \bar{\mathbf{U}} \succeq \mathbf{0}} \log \frac{|\mathbf{\Sigma} + \mathbf{h} w \mathbf{h}^H|}{|\mathbf{\Sigma}|} \\ &- w - \text{tr}(\mathbf{\Omega} \bar{\mathbf{U}}) + \xi \text{tr}(\mathbf{D}) + \bar{\mathbf{p}}^T \boldsymbol{\psi} \end{aligned} \quad (71)$$

where $\bar{\mathbf{p}} = [\bar{\mathbf{p}}^T \quad \bar{\mathbf{I}}^T]^T$ and $\boldsymbol{\psi} = [\boldsymbol{\lambda}^T \quad \boldsymbol{\mu}^T]^T$.

The dual problem is obtained by minimizing $\mathcal{D}(\mathbf{D}, \psi)$ as

$$\min_{\psi \geq 0, \mathbf{D} \succeq 0} \mathcal{D}(\mathbf{D}, \psi). \quad (72)$$

or equivalently as

$$\min_{\psi \geq 0, \mathbf{D} \succeq 0} \max_{w \geq 0, \bar{\mathbf{U}} \succeq 0} \log \frac{|\mathbf{\Sigma} + \mathbf{h}w\mathbf{h}^H|}{|\mathbf{\Sigma}|} - w - \text{tr}(\mathbf{\Omega}\bar{\mathbf{U}}) + \xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi. \quad (73)$$

To obtain an optimal point for the minimax problem in (73), we introduce another optimization variable $\varphi \geq 0$. Then, (73) can be rewritten as

$$\min_{\psi \geq 0, \mathbf{D} \succeq 0} \max_{\varphi \geq 0, w \geq 0, \bar{\mathbf{U}} \succeq 0} \log \frac{|\mathbf{\Sigma} + \mathbf{h}w\mathbf{h}^H|}{|\mathbf{\Sigma}|} - \varphi P + \xi \text{tr}(\mathbf{D}) + \mathbf{p}^T \psi \quad (74)$$

s. t. $w + \text{tr}(\mathbf{\Omega}\bar{\mathbf{U}}) \leq \varphi P.$

It is easy to see that (74) is equivalent to (73) since the inequality must hold with equality at optimality; otherwise, we can scale down φ to achieve a strictly larger objective. Next we make a change of variables as

$$\tilde{w} = w/\varphi, \tilde{\psi} = \psi/\varphi, \tilde{\mathbf{D}} = \mathbf{D}/\varphi, \tilde{\mathbf{U}} = \bar{\mathbf{U}}/\varphi. \quad (75)$$

We now consider \tilde{w} , $\tilde{\psi}$, $\tilde{\mathbf{D}}$, and $\tilde{\mathbf{U}}$ as the new optimization variables. Then, (74) can be equivalently expressed as

$$\min_{\tilde{\psi} \geq 0, \tilde{\mathbf{D}} \succeq 0} \max_{\varphi \geq 0, \tilde{w} \geq 0, \tilde{\mathbf{U}} \succeq 0} \log \frac{|\tilde{\mathbf{\Sigma}} + \mathbf{h}\tilde{w}\mathbf{h}^H|}{|\tilde{\mathbf{\Sigma}}|} + \varphi(\xi \text{tr}(\tilde{\mathbf{D}}) + \mathbf{p}^T \tilde{\psi} - P) \quad (76)$$

s. t. $\tilde{w} + \text{tr}(\mathbf{\Omega}\tilde{\mathbf{U}}) \leq P$

where $\tilde{\mathbf{\Sigma}} = \mathbf{\Sigma}/\varphi$. It is easy to see that the optimal dual variable φ^* can be obtained by considering the minimization of (76) over $\tilde{\psi}$ and $\tilde{\mathbf{D}}$. Hence, (76) is the dual of the following problem:

$$\min_{\tilde{\psi} \geq 0, \tilde{\mathbf{D}} \succeq 0} \max_{\varphi \geq 0, \tilde{w} \geq 0, \tilde{\mathbf{U}} \succeq 0} \log \frac{|\tilde{\mathbf{\Sigma}} + \mathbf{h}\tilde{w}\mathbf{h}^H|}{|\tilde{\mathbf{\Sigma}}|} \quad (77)$$

s. t. $\tilde{w} + \text{tr}(\mathbf{\Omega}\tilde{\mathbf{U}}) \leq P$
 $\xi \text{tr}(\tilde{\mathbf{D}}) + \mathbf{p}^T \tilde{\psi} \leq P.$

Finally, from the derivations in (70), (75), (77) and $\mathbf{W} = \mathbf{\Sigma}^{-1/2} \bar{\mathbf{W}} \mathbf{\Sigma}_k^{-1/2}$, the proof is finalized.

APPENDIX D PROOF OF THEOREM 2

The dual objective of (49) is

$$\begin{aligned} g(\mathbf{v}) &= \inf_{\Gamma_{\text{tol}}} \mathcal{L}(\Gamma_{\text{tol}}, \mathbf{v}) = -\rho^T \mathbf{v} + \inf_{\Gamma_{\text{tol}}} (\tilde{h}(\Gamma_{\text{tol}}) + \mathbf{v}^T \mathbf{a} \Gamma_{\text{tol}}) \\ &= -\rho^T \mathbf{v} - \sup_{\Gamma_{\text{tol}}} ((-\mathbf{a}^T \mathbf{v})^T \Gamma_{\text{tol}} - \tilde{h}(\Gamma_{\text{tol}})) \\ &= -\rho^T \mathbf{v} - \tilde{h}^*(-\mathbf{a}^T \mathbf{v}) \end{aligned} \quad (78)$$

where $\tilde{h}^*(\cdot)$ is the convex conjugate of $\tilde{h}(\cdot)$.

For the convex conjugate $\tilde{h}^*(-\mathbf{a}^T \mathbf{v})$, we use the Legendre Transform to a log function as: $-\log(x) \rightarrow -(1 + \log(-x^*))$ [38]. Thus, the dual problem is

$$g(\mathbf{v}) = -\rho^T \mathbf{v} - 1 - \log(-\mathbf{a}^T \mathbf{v}) \quad (79)$$

and the proof is completed.

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers and the Editor, Prof. N. Kiyavash, for their constructive comments.

REFERENCES

- [1] V.-D. Nguyen, T. Q. Duong, O. A. Dobre, and O.-S. Shin, "Secrecy rate maximization in a cognitive radio network with artificial noise aided for MISO multi-eves," in *Proc. IEEE Inter. Commun. Conf. (ICC)*, Kuala Lumpur, Malaysia, May 2016.
- [2] A. D. Wyner, "The wire-tap channel," *Bell System Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [3] M. Bloch and J. Barros, *Physical-Layer Security*. Cambridge Univ. Press, 2011.
- [4] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 4691–4972, Aug. 2011.
- [5] Y. Pei, Y.-C. Liang, L. Zhang, K. C. Teh, and K. H. Li, "Secure communication over MISO cognitive radio channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1494–1502, Apr. 2010.
- [6] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas-Part I: The MISOME wiretap channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3088–3104, July 2010.
- [7] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas-Part II: The MISOME wiretap channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
- [8] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, June 2008.
- [9] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.
- [10] T. V. Nguyen and H. Shin, "Power allocation and achievable secrecy rates in MISOME wiretap channels," *IEEE Commun. Lett.*, vol. 15, no. 11, pp. 1196–1198, Nov. 2011.
- [11] X. Zhang, X. Zhou, and M. R. McKay, "On the design of artificial-noise-aided secure multi-antenna transmission in slow fading channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2170–2181, June 2013.
- [12] N. Yang, S. Yan, J. Yuan, R. Malaney, R. Subramanian, and I. Land, "Artificial noise: Transmission optimization in multi-input single-output wiretap channels," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1771–1783, May 2015.
- [13] H.-M. Wang, T. Zheng, and X.-G. Xia, "Secure MISO wiretap channels with multiantenna passive eavesdropper: Artificial noise vs. artificial fast fading," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 94–106, Jan. 2015.
- [14] F. Zhu, F. Gao, M. Yao, and H. Zou, "Joint information and jamming beamforming for physical layer security with full duplex base station," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6391–6401, Dec. 2014.
- [15] Q. Li and W. K. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704–2717, May 2013.
- [16] S. Gerbrach, C. Scheunert, and E. A. Jorswieck, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inform. Forensics Security*, vol. 7, no. 2, pp. 704–716, Apr. 2012.
- [17] J. Huang and A. L. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1696–1707, Apr. 2012.
- [18] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.
- [19] F. Zhu and M. Yao, "Improving physical layer security for CRNs using SINR-based cooperative beamforming," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–6, Mar. 2015.
- [20] V.-D. Nguyen, T. Q. Duong, and O.-S. Shin, "Physical layer security for primary system: A symbiotic approach in cooperative cognitive radio networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, Dec. 2015.
- [21] Y. Pei, Y.-C. Liang, L. Zhang, K. C. Teh, and K. H. Li, "Secure communication in multiantenna cognitive radio networks with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1683–1693, Apr. 2011.
- [22] V.-D. Nguyen, T. M. Hoang, and O.-S. Shin, "Secrecy capacity of the primary system in a cognitive radio network," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3834–3843, Aug. 2015.

- [23] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Aug. 2014.
- [24] P. Gopala, L. Lai, and H. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.
- [25] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2735–2751, June 2008.
- [26] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.
- [27] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: an optimized artificial noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.
- [28] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
- [29] L.-N. Tran, M. Juntti, M. Bengtsson, and B. Ottersten, "Beamformer designs for MISO broadcast channels with zero-forcing dirty paper coding," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1173–1185, Mar. 2013.
- [30] K. C. Toh, M. J. Todd, and R. H. Tutuncu, "SDPT3: A Matlab software package for semidefinite programming, version 1.3," *Optimiz. Methods and Softw.*, vol. 11, pp. 545–581, Jan. 1999.
- [31] J. F. Sturm, "Using SeDuMi 1.02: A Matlab toolbox for optimization over symmetric cones," *Optimiz. Methods and Softw.*, vol. 11–12, pp. 625–653, Sept. 1999.
- [32] J. B. Lasserre, "A trace inequality for matrix product," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1500–1501, Aug. 1995.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, UK, 2007.
- [34] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, 1st Ed. Society for Industrial and Applied Mathematics, 2001.
- [35] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd Ed. Society for Industrial and Applied Mathematics, 2002.
- [36] A. M. Nezhad, R. A. Shandiz, and A. E. Jahromi, "A particle swarm-BFGS algorithm for nonlinear programming problems," *Comput. Oper. Res.*, vol. 40, no. 4, pp. 963–972, Apr. 2013.
- [37] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz, "An interior-point method for semidefinite programming," *SIAM J. Optim.*, vol. 6, no. 2, pp. 342–361, 1996.
- [38] W. Fenchel, "On conjugate convex functions," *Canadi. J. Math.*, pp. 73–77, 1949.



Van-Dinh Nguyen received his B.S. degree in Telecommunications from Ho Chi Minh City University of Technology, Vietnam, in 2012, and his M.S. degree in Electronic Engineering from Soongsil University, Seoul, Korea, in 2015. Currently, he is pursuing his Ph.D. degree in Electronic Engineering at Soongsil University, Seoul, Korea. His research interests include wireless communications, information theory, massive MIMO, stochastic geometry, physical layer security, energy harvesting, and cooperative communications, and cognitive radio.



Trung Q. Duong (S'05–M'12–SM'13) received his Ph.D. degree in Telecommunications Systems from Blekinge Institute of Technology (BTH), Sweden in 2012. Since 2013, he has joined Queen's University Belfast, UK as a Lecturer (Assistant Professor). His current research interests include physical layer security, energy-harvesting communications, cognitive relay networks. He is the author or co-author of 190 technical papers published in scientific journals and presented at international conferences.

Dr. Duong currently serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, IET COMMUNICATIONS, WILEY TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES, and ELECTRONICS LETTERS. He has also served as the Guest Editor of the special issue on some major journals including IEEE JOURNAL IN SELECTED AREAS ON COMMUNICATIONS, IET COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS MAGAZINE, IEEE COMMUNICATIONS MAGAZINE, EURASIP JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING, EURASIP JOURNAL ON ADVANCES SIGNAL PROCESSING. He was awarded the Best Paper Award at the IEEE Vehicular Technology Conference (VTC-Spring) in 2013, IEEE International Conference on Communications (ICC) 2014. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2015–2020).



Octavia A. Dobre (M'05–SM'07) is a Professor with the Faculty of Engineering and Applied Science, Memorial University, Canada. Previously, she was with New Jersey Institute of Technology, USA and Polytechnic Institute of Bucharest, Romania. She was a Visiting Professor at Universit  de Bretagne Occidentale, France, and Massachusetts Institute of Technology (MIT), USA (2013). She was the recipient of a Royal Society scholarship at Westminster University, UK (2000), and held a Fulbright fellowship at Stevens Institute of Technology, USA

(2001).

Her research interests include blind signal identification and parameter estimation, physical layer security, cognitive radio, transceiver optimization algorithms, and co-operative communications among others. She has co-authored +180 journal and conference papers in these areas. Her research has been supported by the Natural Sciences and Engineering Research Council of Canada, Mathematics of Information Technology and Complex Systems, Canada Foundation for Innovation, Research and Development Corporation, Atlantic Canada Opportunities Agency, Defence and Research Development Canada, Communications Research Centre Canada, Altera Corporation, Allen Vanguard, DTA Systems, ThinkRF, and Agilent Technologies.

Octavia is the Editor-in-Chief of the IEEE COMMUNICATIONS LETTERS. She was a senior editor, editor and guest editor for diverse prestigious journals. She served as the General Chair of the CWIT 2015, as well as Symposia Co-Chair at various conferences, such as IEEE ICC and Globecom. She is the Chair of the IEEE ComSoc Women in Communications Engineering standing committee and a member-at-large of the Administrative Committee of the IEEE I&M Society. Octavia is a registered Professional Engineer in the province of Newfoundland and Labrador, Canada.



Oh-Soon Shin (S'00–M'10) received his B.S., M.S., and Ph.D. degrees in Electrical Engineering and Computer Science from Seoul National University, Seoul, Korea, in 1998, 2000, and 2004, respectively. From 2004 to 2005, he was with the Division of Engineering and Applied Sciences, Harvard University, MA, USA, as a Postdoctoral Fellow. From 2006 to 2007, he was a Senior Engineer at Samsung Electronics, Suwon, Korea. In September 2007, he joined the School of Electronic Engineering, Soongsil University, Seoul, Korea, where he is currently an Associate Professor. His research interests include communication theory, wireless communication systems, and signal processing for communication.